Tulsa Law Review

Volume 43
Issue 2 *Daubert, Innocence, and the Future of Forensic Science*

Winter 2007

Toward Evidence-Based Evidence: Supporting Forensic Knowledge Claims in the Post-Daubert Era

Simon A. Cole

TOWARD EVIDENCE-BASED EVIDENCE: SUPPORTING FORENSIC KNOWLEDGE CLAIMS IN THE POST-*DAUBERT* ERA

Simon A. Cole*

As legal scholars begin to take stock of what we might call the "*Daubert* regime," the treatment of expert evidence in law in the period following the United States Supreme Court's watershed decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* in 1993,[1] one of the most persistent issues has been a perceived lack of rigor in the application of *Daubert*'s gate keeping requirement to forensic evidence. Scholars have observed that, at least when it comes to civil law, *Daubert* and its progeny decisions did not have the liberalizing effect on the admission of evidence that early readers of the opinion thought it might.[2] To the contrary, some scholars have described an "exclusionary ethos" surrounding the *Daubert* regime.[3] Scholars who focus on criminal law, however, have detected the opposite situation; they have decried the weakness with which *Daubert* has been applied in criminal law, particularly in regard to forensic evidence.[4] What accounts for this disparity? Professor Risinger's comparative examination of outcomes of admissibility decisions across criminal and civil law is disconcertingly consistent with what a legal realist would predict: Trial judges operating under the *Daubert* regime are extremely unlikely to exclude expert evidence proffered by the government in criminal cases, and in civil cases they are far more likely to exclude expert evidence proffered by plaintiffs than by defendants.[5]

Indeed, case law strongly supports the conclusion that the exclusion under the

---

* Associate Professor of Criminology, Law & Society, University of California, Irvine; Ph.D. (Science & Technology Studies), Cornell University; A.B., Princeton University. This project was funded in part by the National Science Foundation (Award #SES-0347305). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. The author consulted, and in some cases testified or submitted affidavits, for the defendants in the following cases that are discussed in this article: *United States v. Mitchell*, *United States v. Llera Plaza*, *Cole*, and *Rose*.

1. 509 U.S. 579 (1993); *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999); *Gen. Elec. Co. v. Joiner*, 522 U.S. 136 (1997).

2. *See generally* Margaret A. Berger, *What Has a Decade of* Daubert *Wrought?* 95 Am. J. Pub. Health 59 (2005).

3. Gary Edmond & David Mercer, Daubert *and the Exclusionary Ethos: The Convergence of Corporate and Judicial Attitudes toward the Admissibility of Expert Evidence in Tort Litigation*, 26 L. & Policy 231 (2004).

4. Paul Giannelli, *Forensic Science*, 34 J.L., Med. & Ethics 310, 310 (2006).

5. D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?* 64 Alb. L. Rev. 99 (2000).

Daubert/Kumho standard for forensic evidence proffered by the government remains rare indeed, despite some recent exclusions.[6] Moreover, the courts' "inclusionary ethos" seems remarkably inconsistent with the trend of legal scholarship, which tends to evince skepticism that many forensic techniques satisfy the requirements of *Daubert* and *Kumho*. This pattern is particularly clear in regard to fingerprint evidence in which numerous courts have found the evidence admissible, despite a body of legal literature that overwhelmingly states the opposite.[7] In this article, I will suggest that one possible explanation for these apparent disparities may be a lack of clarity about the nature and purpose of the *Daubert* inquiry[8] itself. I propose the metaphor "evidence-based evidence," drawing on the culturally resonant prefix "evidence-based," as a way of perhaps clarifying the goal that a *Daubert* inquiry seeks to achieve.

I. *DAUBERT* AND THE ADMISSIBILITY OF EVIDENCE

What is evidence? Evidence might be conceived as the building blocks that contribute the law's edifice of proof. Law is, at least in part, a truth-seeking institution, and it is through the hearing and consideration of evidence that law purports to discover truth. More than that, it is at least to some extent, *only* through evidence that proof may be reached—or fact demonstrated—in law. In this sense, law is not different from other disciplines that claim to be founded on reason, from science to journalism, that likewise claim to rely largely or exclusively on evidence to arrive at factual statements and, more broadly, truth itself.[9]

In trials, law evaluates the relative likelihood of competing factual claims by examining and considering the evidence that supports those claims. By evaluating the evidence legal fact-finders can purportedly make an evaluation as to which factual claim appears more likely to be true. The process necessarily entails a regress, however. In weighing the evidence supporting competing factual claims, a fact-finder might simply assume that all evidence is equally trustworthy. Clearly, however, this is not necessarily the case. Therefore, we must ask: How do we evaluate the relative truth of the evidentiary claims themselves? There appear to be three broad approaches to this problem. The first is that this weighing may be done by the fact-finder simultaneous to the hearing of the evidence, as when a fact-finder is expected to evaluate the trustworthiness of a witness based on demeanor, consistency of answers, ability to withstand cross examination, and so on. The second is through rules, such as the rules of evidence. Of course, it is not necessarily imagined that adherence to these rules always increases the trustworthiness of evidence, but it is perhaps supposed that across the entire

6. *See e.g. U.S. v. Green*, 405 F. Supp. 2d 104 (D. Mass. 2005).

7. David L. Faigman, *Modern Scientific Evidnece: The Law and Science of Expert Testimony* (3d ed., West 2007); Giannelli, *supra* n. 4, at 310.

8. In this article I use the term "*Daubert* inquiry," rather than "*Daubert* hearing" because, although the *Daubert* opinion and its progeny mandate a gate keeping responsibility, they do not mandate that evidence informing the gate keeping decision be gathered specifically via hearing. A court might presumably perform an excellent "*Daubert* inquiry" by gathering information about the reliability about an item of evidence without holding a hearing.

9. *See generally* Susan Haack, *Defending Science—Within Reason: Between Scientism and Cynicism* (Prometheus Bks. 2003).

universe of cases trustworthiness is enhanced. Alternately, such rules may not enhance trustworthiness at all, but may serve other important legal principles.

A third approach descends into the regress, as it were. Rather than merely evaluating the trustworthiness of evidence as it is heard, or adhering to established rules, the tribunal *adjudicates* the trustworthiness of an item of evidence, much as it would adjudicate any other factual claim. Such a process is obviously labor intensive—were it applied to every item of evidence, it would require halting the judicial inquiry each time a new item of evidence was produced in order to adjudicate the trustworthiness of that item of evidence. Therefore, this approach must be used sparingly. This is the approach sometimes adopted with regard to expert evidence whose trustworthiness may plausibly be challenged under the legal regime associated with the *Daubert* decision and its progeny cases.[10]

Given the "regress-ive" nature of a *Daubert* inquiry—that it is, in effect, a trial within a trial—a curious phenomenon emerges. The adjudication of the trustworthiness of an item of evidence must necessarily rely on—what else?—evidence. Thus, a *Daubert* inquiry seeks evidence about evidence, and the very word "evidence" takes on dual, confusing meanings.

It should be noted that, while the regress is potentially infinite, *Daubert* artificially halts the regress after one step. In principle, the regress could continue, like nested Russian dolls. Courts do not, as far as I know, have *Daubert* hearings within *Daubert* hearings. Presumably, the justification is merely pragmatic: There must be some limits to the courts' inquiries into the epistemological foundations of the evidence they use.

The question of fact to be decided in a *Daubert* inquiry is the trustworthiness—or, to use the Court's term—the "reliability" of the item of evidence that will then be inserted into the larger trial that envelops the *Daubert* inquiry as evidence in support of some larger question of fact, such as criminal culpability or civil liability. The court is asked to decide the question of the reliability of evidence by reference to evidence—evidence about the reliability of the evidence. *Daubert* holds that the court must find that the evidence supporting the claim that the evidence is reliable before allowing the evidence to be inserted into the enveloping trial.[11]

Daubert, in short, insists that one form of evidence, expert evidence, be found reliable before being used in a trial. It requires the court to, as Professor Haack puts it, "[t]hink before you think!"[12] But, this finding of reliability, in turn, must be based on evidence. One way of conceptualizing the *Daubert* approach to expert evidence, a way that has not to my knowledge hitherto been proposed, is by reference to a movement that arose nearly contemporaneously with *Daubert* itself: The rise of "evidence-based medicine" (EBM) and the ensuing extension of this metaphor to other areas, including

10. As far as I know, *Daubert* hearings are the only legal procedures that adopt this approach. Courts do not hold evidentiary hearings to assess the reliability of, say, eyewitness evidence. Professor Natapoff has, however, proposed such an approach to informant testimony. Alexandra Natapoff, *Beyond Unreliable: How Snitches Contribute to Wrongful Convictions*, 37 Golden Gate U. L. Rev. 107, 109, 112–15 (2006).

11. *Daubert*, 509 U.S. at 579–98.

12. Susan Haack, *An Epistemologist in the Bramble-Bush: At the Supreme Court with Mr. Joiner*, 26 J. Health, Pol., Policy & L. 217, 217 (2001) (citation omitted).

"evidence-based policy,"[13] "evidence-based corrections,"[14] "evidence-based policing,"[15] "evidence-based crime prevention,"[16] "evidence-based justice,"[17] "evidence-based sociology,"[18] and even the "evidence-based society."[19] In this vein, *Daubert* might usefully be conceived of as a demand in law for "evidence-based evidence," or, more precisely, "evidence-based expert evidence." In this essay, I will briefly explain the concept of evidence-based medicine and its spread to other policy arenas. I will then explore how the "evidence-based" metaphor might apply to law. Next, I will explore the applicability of the notion of "evidence-based evidence" to one application of *Daubert* that has proved particularly vexing for courts: Admissibility challenges to latent print (fingerprint) evidence under *Daubert.*

II. EVIDENCE-BASED MEDICINE AND ITS EXTENSIONS TO POLICY

The conceptual origins of what is today called "evidence-based medicine" (EBM) have been variously traced as far back as Avicenna[20] or nineteenth century France.[21] As a contemporary social movement, however, EBM is frequently traced to the work of physician Archie Cochrane and specifically to a 1972 publication, *Effectiveness and Efficiency.*[22] Cochrane called for demanding evaluations of the efficacy of treatments where possible, rather than ordering treatments based on clinical knowledge. Although Cochrane proposed the essential idea of EBM, the term itself was not used in the literature until 1991.[23] It is interesting to note that this was only two years before the *Daubert* opinion, and yet the connection between EBM and *Daubert*, the way in which *Daubert* and EBM may be read as asking nearly precisely the same question, does not appear to have been explicitly explored.[24]

Today, EBM is defined as "the conscientious, explicit, and judicious use of current

13. The term is attributed to Adrian F.M. Smith by Wikipedia. Wikipedia, *Evidence-Based Policy*, http://en.wikipedia.org/wiki/Evidence-based_policy (accessed Apr. 3, 2008). However, Smith does not use the term in his article on the "evidence-based society." Adrian F.M. Smith, *Mad Cows and Ecstasy: Chance and Choice in an Evidence-Based Society*, 159 J. Royal Statist. Soc. Series A 367, 367 (1996). *But see* Ray Pawson, *Evidence-Based Policy: The Promise of 'Realist Synthesis'*, 8 Evaluation 340, 340 (2002).

14. Doris MacKenzie, *Evidence-Based Corrections: Identifying What Works*, 46 Crime & Delinquency 457, 457 (2000).

15. Lawrence W. Sherman, *Evidence-Based Policing: Social Organization of Information for Social Control*, in *Crime and Social Organization* 217 (Elin Waring & David Weisburd eds., Transaction 2002).

16. Lawrence W. Sherman, *Evidence-Based Crime Prevention* (Routledge 2002).

17. Lawrence W. Sherman & Heather Strang, *Evidence-Based Justice*, 365 Lancet 469, 469 (2005).

18. Stefan Timmermans & Emily S. Kolker, *Evidence-Based Medicine and the Reconfiguration of Medical Knowledge*, 45 J. Health & Soc. Behavior 177, 177 (2004).

19. Smith, *supra* n. 13, at 367.

20. Walter J. Daly & D. Craig Brater, *Medieval Contributions to the Search for Truth in Clinical Medicine*, 43 Perspectives Biology & Med. 530, 531 (2000).

21. David L. Sackett et al., *Evidence Based Medicine: What It Is and What It Isn't*, 312 British Med. J. 71 (1996).

22. Archie Cochrane, *Effectiveness and Efficiency: Random Reflections on Health Services* (RMS Press 1972); Stefan Timmermans & Marc Berg, *The Gold Standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care* (Temple U. Press 2003); R. E. Ashcroft, *Current Epistemological Problems in Evidence Based Medicine*, 30 J. Med. Ethics 131, 131 (2004).

23. G. H. Guyatt, *Evidence-Based Medicine*, 114 Annals Internal Med. A (1991).

24. Though the two phenomena have been discussed together, as in John M. Eisenberg, *What Does Evidence Mean? Can the Law and Medicine Be Reconciled?* 26 J. Health, Pol., Policy & L. 369, 369–72 (2001).

best evidence in making decisions about the care of individual patients."[25] The term has an almost comically self-evident sound to it; lay persons might wonder whether medical decisions were *not* based on evidence prior to 1991 or 1972? (This may also explain the term's rhetorical power: Who could be against "evidence-based medicine"?)[26] In fact, however, medicine has long relied on and heavily valued clinical experience in making decisions about diagnoses and treatments. The rise of EBM was spurred by a number of disturbing findings. First, it was noted that many treatments were widely applied without any convincing evidence of their efficacy.[27] Some of these treatments were later found to be ineffective—that is, to do more harm than good.[28] Second, wide discrepancies in the ordering of tests and treatments were found across different geographic areas of practice.[29] The implication was that medical decisions were likely being governed by local practice rather than by, say, published findings in international medical journals (which would presumably produce like practices in different geographic reasons). By emphasizing "evidence," proponents of EBM sought, when possible, to substitute the results of empirical studies, or what promoters of EBM tend to call "external evidence,"[30] for clinical experience. By "external evidence," EBM meant, most of all, the randomized clinical trial (RCT). Ideally, when possible, meta-analysis (one of the forms of evidence at issue in *Daubert* itself, incidentally)[31] could be used to aggregate the results of evidence produced by all credible RCTs.

EBM is not without controversy. Critics have charged that EBM devalues clinical expertise that may be capable of decision making that applies to individual patients, rather than statistical populations.[32] As a sociologist of science, I also recognize that EBM is a discourse, like any other. A turn to EBM, rather than neatly resolving all problems of medical knowledge, may be expected to generate further debates about what counts as "evidence" and what claims should be considered "evidence-based."[33] It is also well recognized that basing medical decisions on aggregated RCTs is an ideal which cannot be achieved for all medical treatments. For many treatments, RCTs may be impossible, unethical, or unfeasibly costly. In response, proponents of EBM have emphasized that clinical experience retains a role in EBM in conjunction with evidence derived from studies. EBM proposes a "hierarchy of evidence" with RCTs at the apex and clinical experience at the base. At its most innocuous, EBM is merely insisting that evidence from controlled studies, when available, not be ignored in favor of clinical

25. Sackett et al., *supra* n. 21, at 71.

26. Mark A. Peterson, *Evidence: Its Meanings in Health Care and in Law*, 26 J. Health, Pol., Policy & L. 191, 191 (2001) ("Evidence is the apple pie of analysis and decision making. Who could be against it?").

27. Timmermans & Kolker, *supra* n. 18, at 177.

28. *Id.* at 182.

29. *Id.* at 188.

30. Sackett et al., *supra* n. 21, at 71.

31. *Daubert*, 509 U.S. at 582–84.

32. Ashcroft, *supra* n. 22, at 134.

33. For sociologically informed and critical discussions of EBM, see Dawn Freshwater & Gary Rolfe, *Deconstructing Evidence-Based Practice* (Routledge 2004); Timmermans & Berg, *supra* n. 22, at 178; Mark Avis & Dawn Freshwater, *Evidence for Practice, Epistemology, and Critical Reflection*, 7 Nursing Phil. 216, 216–17 (2006); Alberto Cambrosio et al., *Regulatory Objectivity and the Generation and Management of Evidence in Medicine*, 63 Soc. Sci. & Med. 189, 197 (2006).

assessments.[34]

Nonetheless, EBM has been enormously influential in health care, and it has proven a powerful metaphor suitable for importation into other areas of technical decision making. In a 1996 lecture, statistician Adrian Smith broadened the metaphor to endorse the notion of an "evidence-based society."[35] Though Smith did not use the term, his lecture has been credited with giving rise to an "evidence-based policy" movement.[36] Smith specifically mentioned penal policy as a potential target of his proposed call for demanding that evidence govern policy, and, indeed, criminologists have developed, "evidence-based policing,"[37] "evidence-based justice,"[38] and "evidence-based crime prevention."[39] In 2005, my own academic department founded a Center for Evidence-Based Corrections.[40] Again, the seemingly comical self-evident nature of the term—was correctional policy not based on evidence prior to 2005?—contained more than a small grain of truth: Numerous commentators have noted that correctional policy is often governed more by politics than by criminological evidence.[41]

Interestingly, Smith also explicitly mentioned law in his 1996 lecture. Because "[f]or many people, the word evidence conjures up an immediate association with the law," Smith found it "somewhat paradoxical . . . that the procedures and protocols of UK law-courts seem so much at odds with the kinds of disciplined scientific reasoning that many of us would see as essential in an evidence-based society."[42] Understood in this way, the *Daubert* regime might be seen as both an effort and opportunity for law to join the "evidence-based society."

III. APPLYING THE EVIDENCE-BASED METAPHOR TO LAW

How would the "evidence-based" metaphor map onto law? At first glance, it would appear that a criminal trial is already the equivalent of EBM. A fact-finder draws upon the best available evidence to determine the truth of competing factual claims. The fact-finder presumably privileges evidence derived from empirical studies, but, like a physician, will draw on clinical and even anecdotal knowledge to supplement information from such studies, or even substitute for it if no such studies are available. This mapping seems appropriate if one conceives EBM in terms of an individual medical practitioner making decisions for an individual patient. The physician draws upon the

34. Ashcroft, *supra* n. 22, at 132, 134 ("It should be obvious . . . that EBM is not designed to be a comprehensive account of medical knowledge but only an account of that part of medical knowledge which is propositional.").

35. Smith, *supra* n. 13, at 367.

36. Wikipedia, *supra* n. 13; Pawson, *supra* n. 13, at 340.

37. Sherman, *supra* n. 15.

38. Sherman & Strang, *supra* n. 17, at 469.

39. Sherman, *supra* n. 16.

40. U. Cal., *Center for Evidence-Based Corrections*, http://ucicorrections.seweb.uci.edu/ (accessed Apr. 3, 2008).

41. *See e.g.* Franklin E. Zimring et al., *Punishment and Democracy: Three Strikes and You're Out in California* (Oxford 2001); Elliott Currie, *Against Marginality: Arguments for a Public Criminology*, 11 Theoretical Criminology 175, 175 (2007).

42. Smith, *supra* n. 13, at 370. Interestingly, the case Smith points to as an example of the law's indifference to statistical evidence is one that my co-authors and I discuss as well in Michael Lynch et al., *Truth Machine: The Contentious History of DNA Fingerprinting* (forthcoming).

best available evidence, to make the best possible decision about the particular "case."

One can readily appreciate that concepts derived from EBM might be useful in the process. In particular, EBM's hierarchy of evidence might prove useful in helping fact-finders weigh the relative value of different types of evidence. Convincing fact-finders to treat evidence derived from RCTs as stronger than evidence derived from clinical experience might prove quite useful. In addition, EBM has endorsed Bayesian methods of weighing evidence,[43] and many of the most sophisticated legal evidence scholars likewise advocate Bayesian approaches for assessing the weight of evidence.[44] However, in this essay I am concerned with the application of EBM not to the problem of weighing evidence, but to its admissibility.

In addition to weighing evidence at the level of individual decision making, EBM also involves another layer, in which information is distilled from the compendium of knowledge to produce the "evidence" that individual medical decisions are supposed to be based upon. EBM does not envision each individual physician locating, reading, analyzing, and meta-analyzing all the available scientific studies concerning a particular treatment each time they consider applying that treatment to a patient. Such a procedure would obviously be impractical, especially in light of studies showing that many medical practitioners report spending less than an hour per week reading journals.[45] Instead, EBM envisions that the process of evaluating the state of knowledge concerning a particular treatment would be done by some centralized body, rather than by each individual practitioner on an ad hoc basis. One of the most common ways in which this "centralizing" or "standardizing" process is achieved is through the production of "clinical practice guidelines," standardized statements about the efficacy or appropriateness of particular treatments in particular situations.[46] It is very often these guidelines that are then supposed to constitute the "evidence" upon which medical decisions should, according to EBM, be "based."

This process of producing knowledge about the efficacy of certain treatments is, I would suggest, analogous, not to the trial, but to the *Daubert* inquiry, in which the court adjudicates, not a factual dispute, but the reliability of an item of evidence intended to be used in a factual dispute. In EBM, this is done by evaluating the available evidence concerning the treatment's efficacy, and it would seem that in a *Daubert* inquiry the process the Supreme Court envisioned would be much the same: Evaluating the available evidence concerning the reliability of the item of evidence at issue.

IV. APPLYING THE NOTION OF EVIDENCE-BASED EVIDENCE TO FORENSIC EVIDENCE

In this section, I want to suggest that conceptualizing the *Daubert* inquiry as demand for evidence-based evidence may prove helpful in understanding an area in

43. A.S. Elstein, *On the Origins and Development of Evidence-Based Medicine and Medical Decision Making*, 53 Inflammation Research 184 (2004).

44. *See e.g.* Terence Anderson et al., *Analysis of Evidence* (2d ed., Cambridge U. Press 2005); Richard Lempert, *Some Caveats Concerning DNA as Criminal Identification Evidence: With Thanks to the Reverend Bayes*, 13 Cardozo L. Rev. 303 (1991).

45. Sackett et al., *supra* n. 21, at 71.

46. Timmermans & Berg, *supra* n. 22; Arnold J. Rosoff, *Evidence-Based Medicine and the Law: The Courts Confront Clinical Practice Guidelines*, 26 J. Health, Pol., Policy & L. 327 (2001).

which the application of *Daubert* has appeared particularly vexing for courts: forensic evidence.

Daubert was a civil case, and it is possible that its impact on evidence used in criminal law was unanticipated.[47] Anticipated or not, it is now well known that numerous areas of forensic evidence have been subjected to vigorous admissibility challenges in the wake of *Daubert.* What has been startling about these challenges is that they have taken on some of the most widely trusted forms of evidence in the American justice system. "The paramount example" in this regard has been latent print evidence.[48] Although it is true that the vast majority of such challenges have been unsuccessful at rendering the evidence inadmissible, and some courts have characterized the challenges as laughable on their face,[49] several decisions seem to indicate that the claim that latent print evidence is inadmissible under *Daubert* is at least arguable. For example, in *United States v. Llera Plaza I*, the court restricted the scope of expert testimony about latent print evidence, based on it failing all the *Daubert* factors save "general acceptance."[50] The court reversed the outcome of this ruling ten weeks later, but many of its findings stood, including the finding that latent print evidence failed two of the *Daubert* factors.[51] In *Virgin Islands v. Jacobs*, the court ruled latent print evidence inadmissible under *Daubert* after the government neglected to make any evidentiary showing in response to the defendant's *Daubert* motion.[52] In *United States v. Crisp*, one Fourth Circuit judge wrote that the trial judge's admission of latent print evidence despite failing to comply with *Daubert* was an abuse of discretion.[53] In *Utah v. Quintana*, one Utah Court of Appeals judge wrote that the defendant was entitled to a jury instruction highlighting the fallibility and lack of standards of latent print analysis.[54] In *United States v. Sullivan*, the court found latent print evidence inadmissible despite finding that it was untested and therefore failed the "testing" prong of *Daubert*, arguably the most important prong.[55] In *Commonwealth v. Patterson*, the Supreme Judicial Court of Massachusetts found a particular form of latent print evidence, known as "simultaneous impressions," inadmissible.[56] In *State v. Langill*, a trial court found latent print evidence inadmissible because the expert witness violated the laboratory's protocol with regard to documentation and blind "verification," a ruling that was reversed on appeal. And, in *State v. Rose*, in a *Frye* jurisdiction, a court, for the first time, issued an unqualified prohibition against latent print evidence.[57]

Despite the limited or full restrictions on the evidence in some of the above cases, the fact remains that by and large the vast majority of admissibility challenges to latent

47. *See generally* Michael J. Saks, *Banishing* Ipse Dixit*: The Impact of* Kumho Tire *on Forensic Identification Science*, 57 Wash. & Lee L. Rev. 879 (2000).
48. Berger, *supra* n. 2, at 64.
49. *U.S. v. Havvard*, 117 F. Supp. 2d 848, 849 (S.D. Ind. 2000).
50. *U.S. v. Llera Plaza (Llera Plaza I)*, 179 F. Supp. 2d 492, 516–17 (E.D. Pa. 2002).
51. *U.S. v. Llera Plaza (Llera Plaza II)*, 188 F. Supp. 2d 549, 549 (E. D. Pa. 2002).
52. 2001 WL 1735083 at *5 (D.V.I. Dec. 28, 2001).
53. 324 F.3d 261 (4th Cir. 2003) (Michael, J., dissenting).
54. 103 P.3d 168, 170 (Utah App. 2004) (Thorne, J., concurring).
55. 246 F. Supp. 2d 700, 704 (E.D. Ky. 2003).
56. 840 N.E.2d 12, 32–33 (Mass. 2005).
57. No. K06-0545 (Md. Cir. Balt. Co. 2007) (mem.).

print evidence proffered by the government have been decided in favor of admitting the evidence. One might conclude this because the forensic technique is strongly supported by the available evidence. There is, however, the curious fact that the courts and legal scholars appear to part company on this issue. Although almost all courts find latent print evidence admissible, almost all the legal scholarship finds it inadmissible.[58] What can explain this discrepancy? I would suggest that the explanation is that latent print evidence does not constitute evidence-based evidence.

V. LATENT PRINT IDENTIFICATION: IS IT EVIDENCE-BASED EVIDENCE?

Latent print individualization is a forensic assay by which a trained analyst, commonly known as a "latent print examiner" (LPE), seeks to determine whether a print

58. *See* David Faigman et al., *Modern Scientific Evidence: The Law and Science of Expert Testimony* § 27-2.3.1, 386 (2d ed. West 2002) [hereinafter Faigman, *Modern Scientific*] ("Woe to fingerprint practice were such [*Daubert* admissibility] criteria applied."); Jane Campbell Moriarty, *Psychological and Scientific Evidence in Criminal Trials* §12:15 (Clark Boardman Callaghan 2004) ("The assumption of the validity of fingerprinting rests upon law, rather than science."); Nathan Benedict, *Fingerprints and the* Daubert *Standard for Admission of Scientific Evidence: Why Fingerprints Fail and a Proposed Remedy*, 46 Ariz. L. Rev. 519, 538 (2004) ("[J]udges have generally relied on their instincts and the long history of judicial acceptance of fingerprint evidence to admit it without serious consideration of the science behind it."); Berger, *supra* n. 2, at 64 ("Clearly . . . the courts are not applying *Daubert* stringently in the criminal context. The paramount example is fingerprint evidence that has never been validated."); Simon A. Cole, *Grandfathering Evidence: Fingerprint Admissibility Rulings from* Jennings *to* Llera Plaza *and Back Again*, 41 Am. Crim. L. Rev. 1189, 1215 (2004) ("It is clear that no studies exist that measure the accuracy of fingerprint examiners when they make conclusions of identification."); Robert Epstein, *Fingerprints Meet* Daubert*: The Myth of Fingerprint "Science" is Revealed*, 75 So. Cal. L. Rev. 605, 657 (2002) ("Having considered the various indicators of reliability set forth by the Supreme Court in *Daubert*, it is evident that at the present time, latent fingerprint identifications do not constitute reliable evidence."); *Evidence—Fingerprint Experts—Seventh Circuit Upholds the Reliability of Expert Testimony Regarding the Source of a Latent Fingerprint*, 115 Harv. L. Rev. 2349, 2352 (2002) ("Fingerprint expert testimony does not survive application of the *Daubert* factors . . . ."); David L. Faigman, *Is Science Different for Lawyers?* 297 Sci. 339, 340 (2002) (fingerprinting has "not been seriously tested"); Paul Giannelli, *Fingerprints Challenged!* 17 Crim. Just. 33, 35 (2002) ("In its interpretation of *Daubert*, *Plaza I* is a well-written opinion. *Havvard* is not."); David H. Kaye, *The Nonscience of Fingerprinting:* United States v. Llera Plaza, 21 QLR 1073, 1087 (2003) ("As *Llera-Plaza I* so clearly reveals, this [the evidence advanced in support of the admissibility of latent fingerprint individualization] does not satisfy *Daubert*."); Tara Marie La Morte, *Sleeping Gatekeepers:* United States v. Llera Plaza *and the Unreliability of Forensic Fingerprinting Evidence under* Daubert, 14 Alb. L.J. Sci. & Tech. 171, 173 (2003) (discussing "strong indications that the fingerprinting field should not survive a rigorous *Daubert* analysis"); Tamara F. Lawson, *Can Fingerprints Lie? Re-weighing Fingerprint Evidence in Criminal Jury Trials*, 31 Am. J. Crim. L. 1, 65 (2003) ("Currently fingerprint analysis is under attack because of the lack of study done on the accuracy of the examiners . . . ."); Jennifer L. Mnookin, *Fingerprint Evidence In An Age of DNA Profiling*, 67 Brook. L. Rev. 13 (2001) ("In the case of fingerprinting, the general rate of error is simply not known."); Jennifer L. Mnookin, *Fingerprints: Not a Gold Standard*, 20 Issues Sci. & Tech. 47 (2003) ("Judge Pollak's first opinion [restricting latent fingerprint individualization testimony] was the better one."); Michael Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 Hastings L.J. 1069, 1106 (1998) ("By conventional scientific standards, any serious search for evidence of the validity of fingerprint identification is going to be disappointing. . . . A vote to admit fingerprints is a rejection of conventional science as the criterion for admission. A vote for science is a vote to exclude fingerprint expert opinions."); Katherine Schwinghammer, *Fingerprint Idenitification: How "The Gold Standard of Evidence" Could Be Worth Its Weight*, 32 Am. L. Crim. L. 265, 266 (2005); Jessica M. Sombat, *Latent Justice:* Daubert*'s Impact on the Evaluation of Fingerprint Identification Testimony*, 70 Ford. L. Rev. 2819, 2825 (2002) ("[T]he result Judge Pollak reached when he excluded expert testimony concerning fingerprints [in *Llera Plaza I*] was fair."); James E. Starrs, *Judicial Control Over Scientific Supermen: Fingerprint Experts and Others Who Exceed the Bounds*, 35 Crim. L. Bulletin 234, 243 (1999) ("Instead of meaning incapable of error, fingerprint identifications are declared to be infallible on account of the uniqueness of fingerprints to each person . . . ."); Sandy L. Zabell, *Fingerprint Evidence*, 13 J.L. & Policy 143, 178 (2005) ("*ACE-V is an acronym, not a methodology*.") (emphasis in original).

of unknown origin, commonly known as a "latent print" but which we will here, adopting Champod et al.'s parlance, call a "mark,"[59] was made by a particular individual (often a suspect, but sometimes a victim or other person). The claim that the targeted individual made the mark is an inference based on the determination that the mark is not excessively inconsistent with an area of what is known as "friction ridge skin" found on that individual's body. "Friction ridge skin" is an anatomical term for the corrugated skin found on primate fingertips, palms, and soles. Examination of these areas show that they are traversed by lines (known as "ridges") and that these ridges often curve, branch, and end abruptly. The result is that these anatomical areas are covered with a complex weave of curving, branching, and connecting ridges that bears the appearance of an extremely intricate railroad switching yard.

The inference that the mark came from the targeted individual is typically not made from the friction ridge skin itself, but rather from what is commonly known as an "inked print," but which we will call, again adopting Champod et al.'s terminology, simply a "print," to be distinguished from a "mark." A "print" is a deliberately recorded image of the friction ridge skin. Historically, prints were typically made using ink pressed onto paper; today they are often digitally scanned. In either case, the print is, of course, not an exact replica of the friction ridge skin, but an imperfect two-dimensional representation of a three-dimensional structure.

Thus, the examiner seeks to determine whether the mark is consistent with the print. Since the origin of the print is factually known—because it was taken from an individual in custody or from an individual (such as a cooperating witness) whose identity is otherwise, for practical purposes, unquestioned—if the mark and print are consistent, the examiner infers that the individual who is factually known to be the source of the print is also the source of the mark. Reasoning thusly, the examiner testifies that a particular individual is the source of a mark, which can be enormously powerful testimony in resolving a legal matter.

One obvious question raised by this description of the process is: What is a finding of consistency? It is important to note that latent print identification is not based on a finding of *identicality*—that the mark and the print are identical. In fact, while the (rather unspecific, as the reader of the above several paragraphs will now recognize) truism "no two fingerprints are exactly alike" is well known, latent print examiners hold equally to the truism that no two *prints*, even from the source are exactly alike.[60] And, indeed an examination of marks and prints from cases shows that they are not identical at all. The claim, after all, is not that they are identical, but that they derive from a common source.

In short, latent print analysis generates evidence that a particular area of skin is the source of a particular mark. What sort of evidence would allow a court to assess the reliability of this form of evidence? (It should be noted that this is quite a different question from asking whether or not the evidence is "science," "useful," or generally

59. Christophe Champod et al., *Fingerprints and Other Ridge Skin Impressions* (CRC 2004).

60. Wikipedia, *Fingerprint*, http://en.wikipedia.org/wiki/Fingerprint (accessed Apr. 3, 2008) ("The flexibility of friction ridge skin means that no two finger or palm prints are ever exactly alike (never identical in every detail), even two impressions recorded immediately after each other.").

"good.") The most obvious answer would be some sort of measurement of the accuracy of LPEs' source attributions—a test of their ability to make correct source attributions. How often are LPE source attributions correct and how often incorrect?

Such a measurement cannot be derived from casework because in casework we lack access to "ground truth," knowledge of the true source of any particular mark. Not even the corroboration of a second expert, known in the trade as "verification," or even corroboration by an adversarial expert, hired on behalf of the accused, can provide us with ground truth. Nor does a jury's conclusion that a defendant is guilty beyond a reasonable doubt of a particular criminal offense constitute ground truth that that defendant was, in fact, the source of a mark found at the scene of that offense. Indeed, this is even more so when the jury's conclusion is based, in whole or in part, on an LPE's opinion that the defendant was the source of the mark. Because accuracy cannot be measured through casework, accuracy measurement requires deliberately conducting a simulation in which ground truth can be controlled by the experimenter. The obvious method would be to manufacture marks deliberately so that their true origin is known to the experimenter.

Of course, a sophisticated study of this empirical question would not yield a simple binary answer, such as "95% correct." Rather, a sophisticated study would presumably yield accuracy rates that varied according to certain parameters. The most obvious ones are the quality and quantity of information available in the latent print and the skill level of the examiner, but other parameters might also have an impact of the accuracy of latent print analysis. For example, Professors Denbeaux and Risinger argue that they are essentially willing to assume that identifications made from good-quality impressions of all ten fingers are always correct.[61] They are also willing to assume that this finding of absolute correctness would extend down to some smaller amounts of information, though how much smaller they do not know. Professors Denbeaux and Risinger correctly point out that *Kumho Tire*'s "task at hand" requirement dictates that courts should distinguish inquiries into reliability according to the difficulty of various tasks. That is to say: The question "How accurately can LPEs make source attributions for complete sets of ten prints of good quality?" is quite different from the question "How accurately can LPEs make source attributions for single partial latent prints of marginal quality?" The two questions are quite different and clearly should not yield a single common (or "global," as Professors Denbeaux and Risinger put it) answer.

In this, Professors Denbeaux and Risinger are undoubtedly correct, but the more vexing question is what to do in the current situation in which the proponent of the evidence has not differentiated its claim into appropriate subtasks. In latent print admissibility hearings, the government has put forward a "global" claim: That latent print source attributions are reliable for all items of evidence from which latent print examiners choose to make source attributions. In earlier historical periods, and in some countries still, latent print examiners limited their claims by a number of corresponding

61. Mark P. Denbeaux & D. Michael Risinger, Kumho Tire *and Expert Reliability: How the Question You Ask Gives the Answer You Get*, 34 Seton Hall L. Rev. 15, 68–69 (2004).

ridge characteristics, or "points."[62] That is, the claim was: "latent print identification is accurate for latent prints containing more than twelve [or some other number] ridge characteristics." Today, for most U.S. practitioners and law enforcement agencies, the claim is no longer limited in this fashion. Instead, latent print examiners are expected to report conclusions of identity only for those latent prints for which they believe accuracy is assured. This yields the rather vaguer claim: "latent print identification is accurate for those latent prints which examiners believe are 'identifiable.'" In other words, the proponent of the evidence does not concede the seemingly self-evident notion that there must be a gradation of accuracy according to the amount of information in the object being analyzed.

The situation is further complicated by the fact that no scale exists upon which the amount of information in a mark can be specified. Whereas, in the document examination area that is the primary focus of Professor Denbeaux and Risinger's work, the subdivision of tasks into subtasks is seemingly self-evident, it is not obvious how to subdivide latent print source attribution tasks, especially without a scale with which to measure the amount of information in a mark.

One question in this situation is how a court should respond when presented with a global claim of this sort. It seems to me that a court would have difficulty imposing a differentiation of tasks upon the proponent of the evidence, and the court would simply have to evaluate the evidentiary claim as it is given by the proponent. Another question is how a scholar should respond when presented with a global claim of this sort. Here Professors Denbeaux and Risinger and I part company in that they appear to feel a greater obligation to differentiate tasks and concede the reliability of latent print source attributions at the easier end of the continuum of task difficulty. I tend to think that it is the responsibility of the expert making a knowledge claim to specify their claim and have it evaluated as they specify it. I, therefore, feel less obligated to differentiate latent print examiners' tasks since, in the face of all reason, they make global claims to accuracy for all tasks. One possible rationale for such a stance, is that the expert community should bear a cost for making what, as Professors Denbeaux and Risinger correctly point out, is an excessively global claim.

In any case, an accuracy measurement, preferably gradated according to the amount of information contained in the latent print and perhaps other variables as well, is the sort of evidence about latent print evidence that a court might expect to find. The reason there is currently a legal controversy over the admissibility of latent print evidence, however, is that no such evidence has yet been proffered by the government in response to any challenge to the admissibility of latent print evidence.

This is not the same as saying that no such evidence exists. The accuracy data that does exist is quite poor, but some data from simulations in which ground truth was known does, in fact, exist. One source of such data derives from proficiency tests

62. Christophe Champod, *Edmond Locard—Numerical Standards and 'Probable' Identifications*, 45 J. Forensic Identification 136 (1995); Simon A. Cole, *What Counts for Identity? The Historical Origins of the Methodology of Latent Fingerprint Identification*, 12 Sci. Context 139 (1999); *European Fingerprint Standards*, 28 Fingerprint World 19 (2002) (reporting fingerprint point standards ranging from 8 [Bulgaria] to 16 [Italy, Cyprus, Gibraltar] points, as well as some countries with no set standard).

conducted between 1983 and the present by Collaborative Testing Services (CTS) in conjunction with the American Society of Crime Laboratory Directors.[63] This is not ideal data from which to generate accuracy measurements. First, the proficiency tests were conducted by mail. The amount of time taken to complete the tests and the number of individuals who completed each test are not known. The qualifications of the individuals who completed the tests are not known. The difficulty level of the test items is not known. Finally, the proficiency tests were not "masked." In other words, the test-takers knew that they were taking a test. A masked proficiency test would arguably better replicate the accuracy of actual casework. For all of these reasons, it can be argued that the CTS proficiency tests provide only a very crude accuracy measurement for actual latent print casework. Nonetheless, in the absence of any other data, some researchers have compiled the accuracy rate on CTS tests.[64]

Another source of accuracy data is a study conducted by Wertheim et al., of the accuracy of trainees during instruction in latent print analysis.[65] Again, the data is far from ideal. The examiners were trainees, with varying levels of experience in latent print casework. They were able to choose the difficulty of the prints they undertook to attribute. They were given "hints" by the instructors. The study's authors characterized many apparent errors as "clerical errors."[66] Again, these are good reasons to argue that this study provides only a very crude accuracy measurement for actual latent print casework.

The stereotypical contours of argument in legal battles over expert evidence typically consists of studies being put forward by one party followed by methodological critique of those studies by the opposing party. Actors from both sides of the controversy agree that the proficiency test data cited above suffers from numerous flaws. Were the data to be offered as the "evidence" from which the accuracy of latent print identification should be inferred, it would surely be attacked for those flaws. It is important to note, however, that this is not the nature of the legal battle over the admissibility of latent print evidence. Instead, the government has *not* put forward the above potential sources of accuracy data in defending against admissibility challenges to latent print evidence. Indeed, LPEs have publicly criticized defendants' experts for mentioning these sources of data in such hearings.[67] Further, both sources of data contain disclaimers that essentially inoculate them against being used as sources of

63. Collaborative Testing Servs., Inc., *Latent Prints Examination Report Nos. 9508, 9608, 9708, 9808, 99-516, 01-516, 02-516, 02-517, 03-516* (1995–2003). Summaries or complete reports are on file with the author, reports from 2001–2003 available at http://www.collaborativetesting.com/forensics/report_list.html.

64. Lyn Haber & Ralph Norman Haber, *Error Rates for Human Fingerprint Examiners*, in *Automatic Fingerprint Recognition Systems* 339 (Nalini Ratha & Ruud Bolle eds., Springer 2003); Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. Crim. L. & Criminology 985, 1029–32 (2005).

65. Kasey Wertheim et al., *A Report of Latent Print Examiner Accuracy during Comparison Training Exercises*, 56 J. Forensic Identification 55 (2006).

66. Simon A. Cole, *The Prevalence and Potential Causes of Wrongful Conviction by Fingerprint Evidence*, 37 Golden Gate U. L. Rev. 39, 70–71 (2006); Lyn Haber & Ralph Norman Haber, *Letter Re: A Report of Latent Print Examiner Accuracy during Comparison Training Exercises*, 56 J. Forensic Identification 493 (2006); Kasey Wertheim et al., *Authors' Response to Letter*, 56 J. Forensic Identification 500 (2006).

67. Glenn Langenburg, *Defending against the Critic's Curse*, http://www.clpex.com/Articles/CriticsCurse.htm (Sept. 2002).

accuracy data for latent print analysis. Each CTS Report states:

> [Since it is the laboratory's] option how the samples are to be used (e.g. training exercise, known or blind proficiency testing, research and development of new techniques, etc.), the results compiled in the Summary Report are not intended to be an overview of the quality of work performed in the profession and cannot be interpreted as such.[68]

Similarly, the Wertheim et al. study states "[t]hese data should . . . not be used as a predictor of error or an estimate of reliability for an examiner on the witness stand."[69]

It is important to emphasize, therefore, that, as far as courts engaged in admissibility determination are concerned, the above two sources of accuracy data do not exist. They have never been proffered by the government as evidence of the reliability of latent print evidence. Why this is so can only be known for sure by the prosecutors who have handled the admissibility challenges to latent print evidence and those who have advised them, but some speculation is possible. The explanation cannot be that the government feels that these accuracy rates would result in exclusion of the evidence. Although *Daubert* is vague as to precisely how low the error rate of a proffered technique needs to be in order to render it admissible, it seems unlikely that the relatively low false positive error rates found in these studies are above this threshold. Since the government cannot be concerned about the admissibility of the evidence, it must be concerned about its weight. Given that LPEs apparently believe and often testify that latent print identification is "100% accurate" and that it has an error rate of "zero,"[70] one can see why the government might be concerned about introducing even these high accuracy rates into evidence in an admissibility hearing. Once introduced in an admissibility hearing, they would presumably become fodder for cross-examination. Astonishingly, even data showing very high accuracy would have the effect of downgrading the probative value of the evidence from the current status quo.

The analogy with medicine offers a potential explanation for the absence of accuracy measurements of latent print analysis. As noted above, there are many medical interventions that cannot practically, ethically, or cost-effectively generate success rate measurements. In such cases, even proponents of EBM are satisfied to rely on clinical judgments. Is latent print analysis analogous to one of these areas of medicine? Should latent print analysis simply be treated as a clinical judgment?

Certainly, as I have argued elsewhere drawing on the work of the historian Carlo Ginzburg, one can conceive of latent print analyses as clinical judgments.[71] However, I have suggested that they are, in fact, clinical judgments that have been presented to their consumers as something more accurate and precise.[72] In any case, there is nothing about latent print analysis that makes it like one of those areas of medicine for which is practically or ethically unfeasible to generate success rate measurements. While it seems

68. Collaborative Testing Servs., *supra* n. 64.

69. Wertheim et al., *supra* n. 65, at 81.

70. Cole, *supra* n. 64, at 990.

71. Simon A. Cole, *Jackson Pollock, Judge Pollak, and the Dilemma of Fingerprint Expertise*, in *Expertise in Regulation and Law* 98 (Gary Edmond ed., Ashgate 2004); Carlo Ginzburg, *Morelli, Freud, and Sherlock Holmes: Clues and Scientific Method*, in *The Sign of Three: Dupin, Holmes, Peirce* 81 (Umberto Eco & Thomas A. Sebeok eds., Indiana U. Press 1983).

72. Simon A. Cole, *A Little Art, A Little Science, A Little 'CSI'*, N.Y. Times 31 (Dec. 31, 2006).

reasonable to deem admissible clinical judgments that cannot practically or ethically generate success rate measurements, this exemption would not appear to apply to latent print evidence.

VI. LATENT PRINT EVIDENCE IN TRIAL COURT *DAUBERT* INQUIRIES

The upshot of this, of course, is that, as far as courts are concerned, there is *no* accuracy data for latent print source attributions. In other words, there is *no* evidence, of the sort that practitioners of EBM would consider "evidence," as to the accuracy of this form of evidence. In the absence of conventional accuracy data, what sort of evidence have courts relied on in finding latent print expert testimony admissible? It is not possible to answer this question comprehensively because such determinations are made at the trial court level. Many trial courts make such decisions without issuing written rulings, as did the court in the first such challenge in *United States v. Mitchell*. Even if the trial court does issue a written ruling it may not be published.[73] Below, I will discuss the evidence that trial courts have cited in support of the claim that latent print analysis is reliable. For each evidentiary claim, I will explain why it does not constitute evidence of reliability of latent print analysis. Although there are some appellate court rulings concerning the admissibility of latent print evidence, I will not discuss them here. Instead, I restrict my discussion here to direct reports of trial court rulings. Although some of the appellate court rulings do invoke purported evidence of the reliability of latent print evidence, strictly speaking, the issue before the appellate court is not the reliability of latent print evidence itself, but rather whether the trial court's decision was an abuse of discretion.[74] In addition, the appellate court rulings have already been extensively discussed and critiqued in the legal literature.[75]

A. Evidence of Legal Admission and Use of Latent Print Evidence

In some cases, the court treated the fact that latent print analysis has long been considered admissible, reliable evidence in courts of law, as evidence of the reliability of latent print evidence.[76] But a legal finding of reliability cannot be considered scientific evidence of unreliability, unless the legal finding is based upon some sort of scientific evidence of reliability. A look at the cited cases demonstrates that this is not the case—that the legal findings of reliability are assertions devoid of any reference to a scientific study or data.[77]

---

73. Though latent print examiners have made available some of the unpublished decisions. *Daubert* Links, *Legal Challenge to Fingerprints*, http://onin.com/fp/daubert_links.html (last updated Sept. 15, 2005).

74. *Joiner*, 522 U.S. at 138–39.

75. *See e.g.* Benedict, *supra* n. 59; Cole, *supra* n. 59; Faigman, *supra* n. 7; Lawson, *supra* n. 59; Mnookin, *supra* n. 59; Michael J. Saks, *Reliability Standards: Too High, Too Low, or Just Right? The Legal and Scientific Evaluation of Forensic Science (Especially Fingerprint Expert Testimony)*, 33 Seton Hall L. Rev. 1167 (2003); Schwinghammer, *supra* n. 59; Sombat, *supra* n. 59.

76. *See e.g. U.S. v. Frias*, 2003 WL 352502 at *1 (S.D.N.Y. Feb. 13, 2003); *U.S. v. Cruz-Rivera*, 2002 WL 662128 at *1 (D.P.R. Mar. 27, 2002).

77. Cole, *supra* n. 58, at 1275.

B. *Evidence that Latent Print Identification Has Been Used in Court for around a Century*

In some cases, courts have treated the fact that latent print evidence has been used in criminal cases for around one hundred years as evidence of the reliability of latent print evidence. Best known in this regard is *United States v. Havvard*, in which the Southern District of Indiana stated, "the methods of latent print identification can be and have been tested. They have been tested for roughly 100 years. They have been tested in adversarial proceedings with the highest possible stakes—liberty and sometimes life."[78] This reasoning was later criticized in *United States v. Llera Plaza I*.[79] Such assertions cannot serve as evidence of the reliability of latent print identification because the ground truth in casework is not known. Although we know that latent print evidence was used in a large number of criminal cases over the last hundred years, we do not know the frequency with which it produced correct results. This constitutes evidence that the latent print evidence was used, not evidence that it was used correctly. We know that it does not always produce correct results because there are known cases of misidentification.[80] However, we cannot assume that the cases in which latent print analysis produced erroneous results are limited to those cases in which a misidentification has been exposed, agreed upon by practitioners, and made publicly known. In order to make such an assumption, we would have to further assume that we have perfect exposure mechanisms that detect all errors in latent print analysis. There is little reason to believe this is true. Neither "verification" (confirmation of the result by a second examiner from within the analyzing laboratory) nor defense review have been routine for most of the past century. Moreover, both verification and defense review have been implicated in erroneous identifications.[81] In other words, we are aware of known cases in which both verification and defense review *failed* to detect erroneous results.

Moreover, most known erroneous identifications have been exposed through fortuitous events, not routine procedures.[82] This too, suggests that our exposure mechanisms are far from perfect. Finally, as I have shown elsewhere, believing that actual errors are limited to exposed errors would also require believing that errors have been occurring much more frequently over the last two decades than during the preceding eight decades, and that errors occur far more frequently in homicide cases and other serious crimes.[83] No one has advanced an argument as to why this might be the

78. *Havvard*, 117 F. Supp. 2d at 854. *See also State v. Cole*, 2002 WL 1397452 (Del. Super. June 19, 2002); *U.S. v. George*, 2002 WL 1727334 (N.D. Ill. July 24, 2002); *U.S. v. Merritt*, 2002 WL 31854949 (S.D. Ind. Nov. 4, 2002); *U.S. v. Cheshier*, 2001 WL 849346 (S.D. Ind. June 1, 2001); *U.S. v. Cline*, 188 F. Supp. 2d 1287 (D. Kan. 2002); *U.S. v. Reaux*, 2001 WL 883221 (E.D. La. July 31, 2001); *U.S. v. Joseph*, 2001 WL 562988 (E.D. La. May 22, 2001).

79. 179 F. Supp. 2d 492.

80. Cole, *supra* n. 64, at 985.

81. *Id.*

82. *Id.* at 1020–23.

83. *Id.* at 1018–19.

case.

To be sure, the relatively small number of exposed errors over the past century would seem to place some limits on the actual frequency of misidentification. For example, it is difficult to believe that misidentifications occur in half of all analyses and yet are exposed at such a low rate. With proper caution, therefore, an argument might be laid out that treats the large number of deployments of latent print analysis and the low number of exposed errors as vague and weak evidence that the rate of misidentification is not enormously high. Professor Mnookin, a self-described "fingerprint moderate," has articulated how an argument treating the century of use of latent print evidence as a "natural experiment"[84] might be responsibly made.[85] Were a court to adopt Professor Mnookin's argument, intellectual responsibility would require that it acknowledge, at the same time, that there *must* be an unknown number of cases in which the latent print evidence was erroneous, but we have not become aware of its erroneousness.[86] Notably, however, no court has advanced the sort of carefully formulated argument that Professor Mnookin makes. Instead, courts have made the argument that courts should treat as a correct identification each deployment of latent print analysis that is not exposed as a misidentification.

C. *Testimonial Claims That One Laboratory (the FBI Laboratory) Was Not Aware of Having Rendered any Erroneous Conclusions of Individualization*

In *United States v. Llera Plaza II*, the court's finding of the reliability of latent print source attributions made by the Federal Bureau of Investigation Laboratory relied heavily on the testimonial claim of an employee of that laboratory that the FBI was unaware of having made any erroneous identifications.[87] Since ground truth is not known in casework, practitioners' unawareness of having made errors cannot reasonably be treated as evidence of their not having committed errors.[88] In any case, this claim, even if true at the time, can, of course, no longer be made. Within less than two years, the FBI committed an error that was publicly exposed.[89] It should be noted that this error was apparently publicly exposed only because of a media leak; had it not been publicly exposed, courts might still be relying on this supposed "evidence" to make findings of the reliability of latent print evidence.[90]

---

84. Strictly speaking, Professor Mnookin's characterization of this as a "natural experiment" is not correct because this history of use does not create two categories of data, one of which can be characterized as a control and one of which can be characterized as a treatment.

85. Jennifer L. Mnookin, *The Validity of Latent Fingerprint Identification: Confessions of a Fingerprinting Moderate*, ___ L., Probability & Risk ___ (forthcoming).

86. *See* Simon A. Cole, *'Implicit Testing': Can Casework Validate Forensic Techniques?* 46 Jurimetrics 117, 126 (2006).

87. 188 F. Supp. 2d at 575–76.

88. Cole, *supra* n. 64, at 1210; Kaye, *supra* n. 58, at 1087; La Morte, *supra* n. 58.

89. Robert B. Stacey, *A Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case*, 54 J. Forensic Identification 706 (2004).

90. U.S. Dept. Just. Off. Inspector Gen., *A Review of the FBI's Handling of the Brandon Mayfield Case* (unsealed

D. Latent Print Conclusions Can Be Verified by Other Experts

Some courts supported their conclusions that latent print evidence is reliable by reference to the fact that latent print conclusions were subject to review by additional experts.[91] Making a claim that can be refuted (or "falsified" to use the Popperian terminology adopted by the *Daubert* Court) is not the same as providing evidence of the reliability of the claims one makes. But, in any case, such reviews only support the conclusion that experts' opinions are consistent, not that they are accurate. In fact, of the known cases of misidentification, most of them were actually confirmed by additional examiners, many of them by multiple additional examiners, and some of them even by examiners retained on behalf of the defendants.[92]

E. Summary

None of this evidence, even if taken at face value, addresses the question of the accuracy of latent print individualization. In addition, none of the literature defending latent print individualization offers any evidence concerning the accuracy of latent print individualization.[93] In the absence of any information as to the accuracy of latent print individualization conclusions, an informed, reasonable observer certainly might not "accept" conclusions of individualizations. Indeed, while not all expert knowledge claims necessarily lend themselves to conventional validation through controlled experiments, just as not all medical interventions lend themselves to RCTs, given the nature of the latent print examiners' claim—that they can correctly identify the source of latent print to the exclusion of all other possible sources in the universe—any "rationalist" would demand some sort of empirical measurement of their accuracy rate.[94]

It is important to emphasize that this is not a situation in which adversaries dispute the persuasiveness of competing evidence. Government responses to admissibility challenges to latent print evidence consist of arguments, but they do not produce anything that would be recognized as evidence in any rationalist endeavor, like science, medicine, policy, or journalism. Latent print evidence is not evidence-based evidence.

F. Trial Court Rulings Finding an Absence of Evidence Supporting the Reliability of Latent Print Evidence

A minority of trial court admissibility rulings have acknowledged that latent print evidence is not evidence-based evidence. In *United States v. Sullivan*, the Eastern District of Kentucky, noted "that, while the ACE-V methodology appears to be amenable to testing, such testing has not yet been performed."[95] However, the court found "that

91. *See e.g. U.S. v. Salim*, 189 F. Supp. 2d 93, 101 (S.D.N.Y. 2002); *U.S. v. Martinez-Cintron*, 136 F. Supp. 2d 17, 21 (D.P.R. 2001); *Cheshier*, 2001 WL 849346.

92. Cole, *supra* n. 64, at 1023–24.

93. Simon A. Cole, *Is Fingerprint Identification Valid? Rhetorics of Reliability in Fingerprint Proponents' Discourse*, 28 L. & Policy 109 (2006).

94. Edward J. Imwinkelried, *The Meaning of "Appropriate Validation" in* Daubert v. Merrell Dow Pharmaceuticals, Inc., *Interpreted in Light of the Broader Rationalist Tradition, not the Narrow Scientific Tradition*, 30 Fla. St. U. L. Rev. 735, 759 (2003).

95. 246 F. Supp. 2d at 704.

this concern does not render fingerprint evidence unreliable for the purposes of *Daubert*," reasoning that lack of testing went to the weight, not the admissibility, of the evidence.[96]

Another such decision is *Rose*, the first case mandating a blanket exclusion of latent print evidence.[97] Press attention has focused on the court's discussion of the notorious Mayfield case, in which the FBI committed a misidentification.[98] But the commission of a misidentification, even a high-profile misidentification by the FBI and its ratification by an examiner retained by the defendant, does not logically support exclusion of the evidence. First, misidentifications have been known to the courts since the 1920s. More importantly, no admissibility standard demands an absence of error as a condition of admissibility—such a demand would be absurd. Instead, admission requires evidence of reliability.

Rather than being undone by the Mayfield case, a closer reading of the trial court's opinion would seem to suggest that the government simply did not put forward any evidence supporting the reliability of the latent print source attributions. As the court put it, "the State did not prove in this case that opinion testimony by experts regarding the ACE-V method of latent print identification rests upon a reliable factual foundation."[99] The court noted that,

> While the ACE-V methodology appears amenable to testing, such tests have not been performed. The principles underlying ACE-V, that is the uniqueness and permanence of fingerprints, cannot substitute for testing of ACE-V. There have been *no* studies to establish how likely it is that partial prints taken from a crime scene will be a match for only one set of fingerprints in the world.[100]

In its denial of the State's motion for reconsideration, the court further noted that "the Defendant demonstrated that there are no studies of the ACE-V method to determine the reliability of the methodology."[101]

Crucial in this regard is the issue of the burden of proof in an admissibility hearing. Authorities agree that the burden of proof in an admissibility hearing rests upon the proponent of the evidence.[102] However, the *Rose* opinion was among the few opinions

---

96. *Id.*

97. No. K06-0545 (Md. Cir. Balt. Co. 2007). Maryland is a *Frye*, not a *Daubert*, jurisdiction. Nonetheless, a demand for evidence of reliability of the proffered evidence is crucial in the opinion. For an argument that latent print evidence is likewise inadmissible under *Frye*, as it is under *Daubert*, see Simon A. Cole, *Out of the* Daubert *Fire and into the Fryeing Pan? The Admissibility of Latent Print Evidence in* Frye *Jurisdictions*, ___ Minn. J. L., Sci. & Tech. ___ (forthcoming).

98. Jennifer McMenamin, *Judge Bars Use of Partial Prints in Murder Trial*, Balt. Sun 1A (Oct. 23, 2007) ("In explaining her reasoning in a 32-page decision, the judge leaned heavily on the case of an Oregon lawyer mistakenly linked through fingerprint analysis to the 2004 Madrid train bombings."). On Mayfield, see Cole, *supra* n. 65, at 985–87; U.S. Dept. Just. Off. Inspector Gen., *A Review of the FBI's Handling of the Brandon Mayfield Case* (Mar. 2006); William C. Thompson & Simon A. Cole, *Lessons from the Brandon Mayfield Case*, 29 Champion 42 (Apr. 2005); Steven T. Wax & Christopher J. Schatz, *A Multitude of Errors: The Brandon Mayfield Case*, 26 Champion 6 (Oct. 2004).

99. *Rose*, No. K06-0545, slip op. at 1 (quoted from memorandum decision re reconsideration).

100. *Rose*, No. K06-0545, slip op. at 26 (internal citations omitted) (quoted from memorandum decision).

101. *Rose*, No. K06-0545, slip op. at 17 (quoted from memorandum decision re reconsideration).

102. Saks, *supra* n. 76, at 1173 ("Elementary principles of law place the burden of proof on the proponent of the admission of evidence. Accordingly, *Daubert* places the initial burden of production on the proponent of [illegible] expert witness

in the line of latent print admissibility challenges to acknowledge this. In one such case, *Virgin Islands v. Jacbos*, the court excluded latent print evidence in which the government put forward no evidence whatsoever concerning the reliability of latent print evidence.[103] But, in what is probably the best known such case, *United States v. Llera Plaza II*, the court shifted the burden of proof to the defendant making the absence of evidence concerning the accuracy of latent print evidence count against the *opponent* of the evidence.[104] In *United States v. Mitchell*, the court unabashedly shifted the burden to the opponent of the evidence.[105] In *Rose*, however, the court noted that "the burden is on the proponent of the evidence to prove the reliability" of the evidence.[106] It concluded that "the State did not meet that burden in this case."[107] In its denial of the State's Motion for Reconsideration, the court admitted that it was "surprising . . . to this Court that the State was not able to meet its burden of proof in this case," and stated that "it has been shocking to the community."[108]

VII. CONCLUSION

It is indeed shocking that the government appears unable to muster any evidence of reliability for a technique as venerable as latent print identification. The fact that the government cannot support the claim of reliability does not, of course, necessarily mean that that the technique is highly inaccurate. Perhaps one reason that it is so difficult to muster evidence in support of latent print evidence, however, is that courts have been shielding the government from the demand for evidence of reliability. In the pre-*Daubert* era courts allowed latent print evidence to win admissibility based on the *ipse dixit* of its practitioners.[109] In the post-*Daubert* era, they continued to allow admissibility without demanding what any rationalist enterprise would treat as evidence of reliability. These rulings not only protected the government from generating evidence about the reliability of latent print evidence, but may have actually discouraged the government from generating it.

In the case of forensic evidence the situation is similar to that which obtained in medicine at the time of EBM. That is, there are some "treatments" that had been used for a long period of time on the assumption that they are effective without any evidence that they are, in fact, effective. Similarly, there are some forensic techniques that the

103. 2001 WL 1735083.

104. 188 F. Supp. 2d at 576.

105. 365 F.3d 215, 240 (3d Cir. 2004).

106. *Rose*, No. K06-0545, slip op. at 22 (quoted from memorandum decision).

107. *Id.* at 31.

108. *Rose*, No. K06-0545, slip op. at 12 (quoted from memorandum decision re reconsideration). It is interesting to note that, after the *Rose* court denied the government's Motion for Reconsideration, the government indicted Rose in federal court. Brent Jones, *Man Indicted by Federal Grand Jury in 2006 Killing at Mall*, Balt. Sun, (Apr. 2, 2008). The State's Attorney explicitly stated that the exclusion of the latent print evidence in state court was the reason for indicting in federal court. What is particularly interesting for our purposes here about this exercise in forum shopping is the implicit suggestion that the government hopes that the requirement that trial evidence be evidence-based, that apparently is adhered to in the Maryland courts, will not extend to the federal courts. This leads to the rather surprising implication that the government hopes that the admissibility threshold for expert evidence will be less stringent in federal court than in Maryland state court.

109. Saks & Koehler, *supra* note 88.

criminal justice system has been relying on for a long period of time on the assumption that they are reliable without any actual evidence they are, in fact, reliable. As in the case of medical treatments, we should expect that in some cases our assumptions have been well founded, and in other cases they have not. The legal admissibility problem, however, is easier to solve: *Daubert* demands evidence of reliability; it does not allow for the assumption of reliability. This, I would suggest, is part of the explanation for the vexing nature of admissibility challenges to forensic evidence. Many forensic techniques, however accurate they may actually be, simply lack evidence concerning their accuracy. In this situation, a strict reading of *Daubert* demands exclusion even of evidence that may turn out to be highly accurate, until such time as evidence of its accuracy is amassed.

The history of latent print admissibility challenges serves to illustrate the need for conceptualizing *Daubert* inquiries as demands that evidence used in trials be "evidence-based." If courts take seriously the notion that *Daubert* hearings are trials that demand the production of evidence about the reliability of the evidence that parties propose to use in the enveloping trials, perhaps the American legal system will move a step closer to joining the "evidence-based society."