

1-1-2009

Rounding up the Usual Suspects: A Legal and Logical Analysis of DNA Trawling Cases

David H. Kaye

Follow this and additional works at: <http://scholarship.law.unc.edu/nclr>Part of the [Law Commons](#)

Recommended Citation

David H. Kaye, *Rounding up the Usual Suspects: A Legal and Logical Analysis of DNA Trawling Cases*, 87 N.C. L. REV. 425 (2009).Available at: <http://scholarship.law.unc.edu/nclr/vol87/iss2/3>

This Article is brought to you for free and open access by Carolina Law Scholarship Repository. It has been accepted for inclusion in North Carolina Law Review by an authorized administrator of Carolina Law Scholarship Repository. For more information, please contact law_repository@unc.edu.

ROUNDING UP THE USUAL SUSPECTS: A LEGAL AND LOGICAL ANALYSIS OF DNA TRAWLING CASES*

DAVID H. KAYE**

Courts are beginning to confront a problem that has divided the scientific community—whether identifying a defendant by fishing through a database of DNA types to find a match to a crime-scene sample reduces the significance of a match. For years, the problem seemed academic. Now that the United States has more than six million DNA profiles from convicted offenders and suspects in a national, computer-searchable database, the question is more urgent. Increasingly, individuals are being charged with crimes as a result of a match between their recorded profile and the DNA from a victim or scene of a crime. Some of these defendants have moved to exclude the DNA evidence, arguing that there is no generally accepted scientific opinion on how the probative value of a match from a database trawl should be quantified. Trial courts have ruled both ways. Appellate courts in California and the District of Columbia have rebuffed these challenges, reasoning that the general-acceptance standard for scientific evidence does not apply in this situation. Furthermore, they have held that the jury can be given the usual probability that a randomly selected individual will match a crime-scene sample. Elaborating on earlier, Bayesian statistical analyses of the database-trawl question, this Article defends the admissibility of the random-match probability even when the defendant was not selected randomly. It also considers various alternatives to the introduction of this probability and proposes modifications to the “np rule” for adjusting a random-match probability p by the database size n .

* Copyright © 2009 by David H. Kaye.

** Regents’ Professor, Sandra Day O’Connor College of Law, Professor, School of Life Sciences, and Fellow, Center for the Study of Law, Science and Technology, Arizona State University. This Article is adapted from a forthcoming book, *The Double Helix and the Law of Evidence: Controversies over the Admissibility of Genetic Evidence of Identity* (Harvard Univ. Press). Thanks are due to Ronald Allen, Charles Brenner, James Crow, Philip Dawid, Peter Donnelly, Richard Friedman, John Hartmann, Stephen Kanter, Jay Koeler, Michael Risinger, George Sensabaugh, and Stephen Stigler for comments on a draft of this paper or discussions of the subject. Edward Ungvarsky and Bicka Barlow provided helpful information or trial materials from *People v. Jenkins* and *People v. Puckett*, respectively.

INTRODUCTION	426
I. TRAWLING IN DATABASES: THE CASES	432
A. <i>A Murder on Capitol Hill</i>	433
1. A Battle over the DNA Match	435
2. The Easy Way Out: The p_{rem} and np Rules	436
3. I've Got Your Numbers	443
B. <i>California Dreaming: Johnson and Nelson</i>	446
II. TRAWLING IN DATABASES: PROBATIVE VALUE	453
A. <i>The Dangers of Multiple Hypothesis Testing</i>	454
B. <i>The Bayesian Rejoinder</i>	459
1. The Ingredients	463
2. The Recipes	465
3. The Dishes	467
C. <i>Reconciling Conflicting Frequentist and Bayesian Intuitions</i>	469
D. <i>The Naked Trawl and Other Cases</i>	472
1. The Pure Confirmation Case: $n = 1$	473
2. The Trawl-as-confirmation Case: $n > 1$	474
3. The Trawl-then-confirm-successfully Case	476
4. The Trawl-then-confirm-unsuccessfully Case	477
E. <i>Back to Basics</i>	478
III. BAYES IS BETTER?	482
A. <i>Bayes for the Defense</i>	483
B. <i>Other np-like Rules</i>	488
CONCLUSION	492
APPENDIX	494
A. <i>Derivation of Bayes' Rule</i>	495
B. <i>The Impact of Trawl Evidence</i>	496
1. The Pure Confirmation Case: $n = 1$	497
2. The Trawl-as-confirmation Case: $n > 1$	498
3. The Trawl-then-confirm-successfully Case	500
4. The Trawl-then-confirm-unsuccessfully Case	501

INTRODUCTION

Countries around the world have established databases consisting of the DNA profiles of suspected or convicted offenders.¹

1. England is the world's leader, having started the first national criminal DNA database in 1995. In proportion to the population, the British database is "the largest of any country: 5.2% of the UK population is on the database compared with 0.5% in the USA. . . . By the end of 2005 over 3.4 million DNA profiles were held on the database—the profiles of the majority of the known active offender population." U.K. Home Office,

In the United States, state and federal databases in the FBI's National DNA Index System hold over six million DNA profiles of people who have been convicted, arrested, or detained for various crimes.² These identification databases have helped solve cases that have baffled investigators for decades. In one instance, a fifty-eight-year-old man became a suspect in over twenty-four rapes in three states dating back to 1973 as a result of a coordinated database search.³ More recently, a DNA database match to a highly degraded semen sample culminated in the conviction of "John Puckett, an obese, wheelchair-bound 70-year-old" for murdering Diana Sylvester, a "22-year-old San Francisco nurse [who] had been sexually assaulted and stabbed in the heart" in her San Francisco apartment over thirty years ago.⁴ In addition to such dramatic but relatively rare "cold cases," database trawls have considerable potential to solve common property crimes such as automobile theft and petty burglaries.⁵

When the DNA profile from a crime-scene stain matches a DNA pattern on file, the person identified by this "cold hit" becomes a target of the investigation. A fresh sample will be taken from the suspect to verify the DNA match, and other evidence normally will reinforce the investigatory lead. In some cases, prosecutors will even proceed with almost no evidence beyond the cold hit. In one such case, a San Francisco jury convicted John Davis, already behind bars for robbery and other crimes, of the murder of his neighbor, Barbara Martz, nearly twenty-two years earlier.⁶ The database match was all the jurors had to go on.⁷ The prosecution maintained that this match

The National DNA Database, <http://www.homeoffice.gov.uk/science-research/using-science/dna-database/> (last visited Dec. 22, 2008).

2. FBI, CODIS-NDIS Statistics, <http://www.fbi.gov/hq/lab/codis/clickmap.htm> (last visited Dec. 22, 2008).

3. Julia Preston, *After 32 Years, Clothing Yields DNA Key to Dozens of Rapes*, N.Y. TIMES, Apr. 27, 2005, at A1. He was promptly convicted of the thirty-two-year-old rape. Julia Preston, *After 3 Decades, Guilty Verdict in Rape Case, With Help From DNA*, N.Y. TIMES, Nov. 10, 2005, at B1. Of course, database matches are not limited to high profile crimes. See, e.g., Shaila K. Dewan, *As Police Extend Use of DNA, A Smudge Could Trap a Thief*, N.Y. TIMES, May 26, 2004, at A1; Alison Gendar, *DNA Test Buys Rob Busts*, N.Y. DAILY NEWS, Mar. 29, 2005, http://www.nydailynews.com/archives/news/2005/03/29/2005-03-29_dna_test_buoys_rob_busts.html.

4. Jason Felch & Maura Dolan, *When a Match is Far from a Lock*, L.A. TIMES, May 4, 2008, at A1; *People v. Puckett*, No. 201396 (Super. Ct., S. F. County 2008).

5. Dewan, *supra* note 3; Gendar, *supra* note 3; Charlie Jannetto, *DNA Solves Wider Range of Crimes*, ARIZ. REPUBLIC, July 27, 2008, at B1, available at <http://www.azcentral.com/news/articles/2008/07/27/20080727dna0727.html>.

6. Jaxon Van Derbeken, *DNA Nails Neighbor in 1985 Murder*, S.F. CHRON., Aug. 28, 2007, at B1.

7. *Id.*

was enough for a conviction because the “random-match probability”—the probability that a randomly selected, unrelated individual would match the crime-scene DNA sample—was in the quadrillionths.⁸ After discussing the remote possibility of an unknown, identical twin to the defendant, the jury convicted.⁹ In the *Puckett* trial, a much larger (and hence less impressive) random-match probability of one in 1.1 million proved persuasive.¹⁰

Cases like *Davis* and *Puckett* that emanate from cold hits have been called “trawl”¹¹ cases because “the DNA match itself made the defendant a suspect, and the match was discovered only by searching through a database of previously obtained DNA samples.”¹² These database-trawl cases can be contrasted with traditional “confirmation cases” in which “other evidence has made the defendant a suspect and so warranted testing his DNA.”¹³ In terms of this dichotomy, the statistical and evidentiary issue is whether the fact that the defendant was selected for prosecution by trawling requires an adjustment to the normally tiny random-match probability to reflect the greater probability of a match to at least one suspect in the database. Two committees of the National Academy of Sciences (“NAS”) recommended adjusting for the search strategy. In their influential reports on “DNA Forensic Science” released by the National Research Council (“NRC”)¹⁴ in 1992¹⁵ and 1996,¹⁶ the committees reasoned that a match resulting from a trawl is less impressive than a match in a confirmation case.¹⁷ Their reasoning is based on a

8. *Id.*

9. *Id.*

10. Felch & Dolan, *supra* note 4.

11. David J. Balding, *Errors and Misunderstandings in the Second NRC Report*, 37 JURIMETRICS J. 469, 470 (1997).

12. Peter Donnelly & Richard D. Friedman, *DNA Database Searches and the Legal Consumption of Scientific Evidence*, 97 MICH. L. REV. 931, 932 (1999).

13. *Id.*

14. “The National Research Council (NRC) functions under the auspices of the National Academy of Sciences (NAS), the National Academy of Engineering (NAE), and the Institute of Medicine (IOM).” Welcome to the National Research Council, <http://sites.nationalacademies.org/nrc/index.htm> (last visited Dec. 22, 2008). The NRC seeks “to improve government decision making and public policy, increase public education and understanding, and promote the acquisition and dissemination of knowledge in matters involving science, engineering, technology, and health.” *Id.* Although the organizational structure of the National Academies is not crucial here, I refer to the two DNA committees as NAS committees and their reports as NRC reports.

15. NAT’L RESEARCH COUNCIL COMM. ON DNA TECHNOLOGY IN FORENSIC SCI., *DNA TECHNOLOGY IN FORENSIC SCIENCE* (1992) [hereinafter NRC I].

16. NAT’L RESEARCH COUNCIL COMM. ON DNA FORENSIC SCI.: AN UPDATE, *THE EVALUATION OF FORENSIC DNA EVIDENCE* (1996) [hereinafter NRC II].

17. *See infra* Part II.A.

“frequentist” theory of statistical hypothesis testing, which holds that an investigator can reject a particular hypothesis (the “null hypothesis”) in favor of an alternative hypothesis when the probability of observing the data (assuming that the null hypothesis is true) is sufficiently small.¹⁸ For example, we might toss a coin twenty times to test the null hypothesis that the coin is fair. A large imbalance between heads and tails would count as evidence against this null hypothesis. The probability of so large an imbalance is known as a “*p*-value.” Small *p*-values indicate that the evidence is inconsistent with the null hypothesis—we would not see so large a discrepancy between the numbers of heads and tail very often if the coin were fair.

The null hypothesis emphasized by the committees is that the source of the DNA from the crime scene is someone who is not in the database (and hence that the individual caught in the trawl is innocent). The committees concluded that the random-match probability is not the appropriate *p*-value for testing this null hypothesis. After all, in repeated trawls of innocent databases—those that do not include the source of the crime-scene DNA—a match to *someone* in the database would be a more common event than the random-match probability suggests. As the 1996 committees pointed out, observing twenty heads in a row when flipping a fair coin is astonishing—the *p*-value is about one in a million—but the same observation—twenty heads in a row—is less impressive if one has flipped the coin millions of times.¹⁹ To account for this “multiple testing,”²⁰ the NRC reports described approaches that would inflate the normal random-match probability, making a match in a database-trawl case appear to be less probative than in a confirmation case.²¹

18. See, e.g., DAVID H. KAYE ET AL., *THE NEW WIGMORE: A TREATISE ON EVIDENCE: EXPERT EVIDENCE* § 11 (2004) (surveying theories of statistical reasoning and their application to legal proof). More generally, the probability used to evaluate the null hypothesis pertains to all possible data as extreme or more extreme (relative to what would be expected under the null hypothesis) than what was observed. *Id.* § 11.8.2.

19. NRC II, *supra* note 16, at 134.

20. NRC I, *supra* note 15, at 124.

21. *Id.*

The response has been disputation and litigation.²² The adjustments for trawling have been criticized as addressing the wrong hypothesis and as needlessly conservative, but “[t]he mathematical arguments given by [each] side appear impressive” and “there is [so] much confused writing [on the subject that] it would be very difficult for a court to make a reasoned decision based on a simple assessment of [the] literature recommendations.”²³ Only three published opinions consider whether the prosecution can use the random-match probability for a confirmation case in a trawl case. In *United States v. Jenkins*,²⁴ *People v. Johnson*,²⁵ and *People v. Nelson*,²⁶ defendants argued, to no avail, that the scientific community does not agree on a single statistic to characterize the import of a database trawl, and until it does, even the fact of a match should not be admitted.

This legal argument invokes a specialized rule for scientific evidence that requires a showing of “general acceptance” within the scientific community. Having originated in the 1923 case of *Frye v. United States*,²⁷ this requirement is deeply entrenched in the law of many states.²⁸ It is one form of especially strict scrutiny for scientific, as opposed to other types of, expert evidence.²⁹ Another version of strict scrutiny is found in the celebrated opinion of the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*³⁰ and later codified in an amendment to Federal Rule of Evidence 702. *Daubert* instructs judges to ensure that the scientific evidence is truly “reliable” by examining general acceptance along with a variety of

22. The NRC reports (and the assessment of their recommendations by the FBI's DNA Advisory Board, *Statistical and Population Genetics Issues Affecting the Evaluation of the Frequency of Occurrence of DNA Profiles Calculated From Pertinent Population Database(s)*, FORENSIC SCI. COMMS., July 2000, <http://www.fbi.gov/hq/lab/fsc/backissu/july2000/dnastat.htm>, reprinted in JOHN M. BUTLER, FORENSIC DNA TYPING: BIOLOGY, TECHNOLOGY, AND GENETICS OF STR MARKERS 613 (2d ed. 2005)), are not the last words on the subject in the statistical literature. See Simon Walsh & John Buckleton, *DNA Intelligence Databases*, in FORENSIC DNA EVIDENCE INTERPRETATION 439, 464–65 (John S. Buckleton et al. eds., 2005); Geir Storvik & Thore Egeland, *The DNA Database Search Controversy Revisited: Bridging the Bayesian-Frequentist Gap*, 63 BIOMETRICS 922, 922–23 (2007) (reviewing the literature).

23. Walsh & Buckleton, *supra* note 22, at 464–65.

24. 887 A.2d 1013 (D.C. 2005).

25. 43 Cal. Rptr. 3d 587 (Ct. App. 2006).

26. 48 Cal. Rptr. 3d 399 (Ct. App. 2006), *aff'd*, 185 P.3d 49 (Cal. 2008).

27. 293 F. 1013 (D.C. Cir. 1923) (rejecting primitive lie-detector evidence as not generally accepted).

28. See, e.g., KAYE ET AL., *supra* note 18, § 5.3.

29. *Id.* § 5.1.

30. 509 U.S. 579 (1993).

other indicia of scientific validity.³¹ Yet, the appellate courts in Washington, D.C., and California rejected the argument that the “unadjusted random-match probability” has to be evaluated under the general-acceptance standard of *Frye*, and their reasoning applies with equal force to the scientific-validity standard of *Daubert*. With *Frye* (and *Daubert*) out of the way, these courts quickly upheld the admission of the unadjusted random-match probability in the trawl cases.

This Article argues that *Jenkins*, *Johnson*, and *Nelson* reached the correct result. The dispute over adjustment does not involve the general acceptance (*Frye*) or scientific validity (*Daubert*) of the formulas for estimating a random-match probability or the *p*-value for the null hypothesis that the database is innocent. The real issue is which statistic or statistics are logically relevant (and not unfairly prejudicial) in a trawl case. The analysis of relevance and prejudice that I shall present demonstrates that the random-match probability for a DNA profile need not be adjusted to account for a database trawl. Such an analysis is needed because the opinions do not directly address the frequentist argument that the unadjusted random-match probability is unfair and misleading. As a result, they do not demonstrate that admitting the random-match probability rather than an adjusted probability is the correct answer to the relevance inquiry.³²

This Article fills this gap. Part I examines the three appellate opinions in detail. This Part enunciates the rationales that the courts in *Jenkins*, *Johnson*, and *Nelson* relied on and finds them wanting.

Part II presents a better theory for using unadjusted random-match probabilities than the explanations given in *Jenkins*, *Johnson*, and *Nelson*. This Part attempts to accommodate the conflicting intuitions that have led to the scientific divide. It recognizes that the

31. *Id.* at 593–95. For elaboration and analysis of the two standards as well as a discussion of a third “relevancy-helpfulness” approach used in still other jurisdictions, see KAYE ET AL., *supra* note 18, § 5.3.

32. As two statisticians recently observed:

The discussion about [the] approaches [to quantifying the evidence in a trawl case] now has somewhat died out (at least in scientific journals), with most scientists in the field preferring the [no-adjustment] approach. Still, a satisfactory explanation for the differences between these approaches and full understanding with respect to some of the aspects in the criticisms that have appeared is lacking.

Storvik & Egeland, *supra* note 22, at 922. With respect to the developing case law, this assessment is apt. The opinions do not reflect a “full understanding” and do not provide an entirely “satisfactory explanation” for using the random-match probability.

apparent “statistical significance” of a result can evaporate if the finding is the product of a “search for significance.” Nevertheless, it uses an elementary formula of probability theory, Bayes’ rule, to explain why this concern does not apply to the kind of database searches undertaken in criminal investigations.³³

Part III of the Article considers how defendants in trawl cases might respond to the apparently overwhelming random-match probabilities and what courts can do to avoid unfair prejudice. In a reprise to a legal debate that has stretched across four decades,³⁴ it describes an unusual case in which jurors equipped with court-issued pocket calculators were encouraged to apply Bayes’ rule for themselves. Part III does not defend this particular procedure, but it does maintain that, contrary to recent opinions in England and the United States, the defendant should have the opportunity to use the mathematical logic to minimize the impact of the DNA match in this kind of case. It also shows that if defendants are allowed to introduce the adjusted random-match probability advocated in the 1996 NRC report, the adjustment should be modified in a manner that was overlooked in the report.

I. TRAWLING IN DATABASES: THE CASES

Prosecutors have been regaling jurors with unadjusted random-match probabilities in trawl cases for at least fifteen years. In 1994, Troy Bloom was convicted of rape after emerging as one of five suspects generated by a search of Minnesota’s convicted-offender database. After extending the DNA testing to more loci, the state produced testimony that Bloom matched semen in the victim’s car at all the loci and “that the probability of a random match was 1 in 4.6 million.”³⁵ In *Bloom* and other trawl cases, however, defendants did not challenge the propriety of the random-match probability as a

33. Cf. Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1021–22 (1977) (introducing Bayes’ rule as an analytical tool or “heuristic” device to explore the concept of relevance).

34. The seminal articles are Michael O. Finkelstein & William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489 (1970), and Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971). See generally KAYE ET AL., *supra* note 18, § 12.4.3 (describing posterior probabilities and courts’ approaches to them).

35. *Bloom v. State*, No. C8-97-885 (Minn. Ct. App. Feb. 10, 1998), available at <http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=mn&vol=appunpub\9802\885&invol=1> (unpublished opinion denying post-conviction relief). For earlier proceedings, see *State v. Bloom*, 516 N.W.2d 159 (Minn. 1994); David Chanen, *Court of Appeals Upholds '94 Rape Conviction that Used DNA Analysis*, STAR TRIB. (Minneapolis-St. Paul), Feb. 1, 1996, at B3.

measure of the significance of a match. The first concerted challenge to the use of unadjusted probabilities came some ten years later, around 2005. In *United States v. Jenkins*,³⁶ a trial court in Washington, D.C., excluded the random-match probability, but the jurisdiction's highest court overturned this ruling. As defendants continued to press their objection, appellate courts in California issued opinions with different theories as to why the random-match probability is admissible. In *People v. Nelson*,³⁷ California's Supreme Court, citing *Jenkins*, unanimously endorsed one of these theories. This Part describes these developments and suggests that the judicial analyses in this wave of cases are too shallow to dispose of the objection to unadjusted probabilities. We begin with *Jenkins*.

A. *A Murder on Capitol Hill*

Dennis Dolinger was a community activist known for establishing a neighborhood patrol to thwart drug dealers.³⁸ He lived on Capitol Hill in the southeastern section of the District.³⁹ On June 4, 1999, police found his body in the basement of his Potomac Avenue home.⁴⁰ He had been stabbed repeatedly in the head.⁴¹ His wallet, a diamond ring and a gold chain were missing.⁴² Blood was on clothing in the basement and in a room on the second floor.⁴³ More blood drops led from the basement, throughout the house, and out to the yard and sidewalk.⁴⁴ The case seemed to be solved when Stephen Watson, a resident of nearby Alexandria, Virginia, was discovered with Dolinger's credit card and other personal items.⁴⁵ Watson remained in jail for several months awaiting trial,⁴⁶ but the case against him fell apart after the D.C. police sent samples of the trail of blood to the FBI. "[Thirty] or so" of these samples did not match Dolinger's DNA.⁴⁷ Furthermore, they had the same one-person profile, indicating that they all came from a single assailant who must

36. 887 A.2d 1013 (D.C. 2005).

37. 185 P.3d 49 (Cal. 2008).

38. Eric M. Weiss, *Va. Man Convicted In Activist's Killing*, WASH. POST, June 28, 2006, at B9.

39. *Id.*

40. *Jenkins*, 887 A.2d at 1016.

41. *Id.*

42. *Id.*

43. *Id.*

44. *Id.*

45. *Id.*

46. Weiss, *supra* note 38.

47. Transcript of Record at 3, *Jenkins*, 887 A.2d 1013 (No. F320-002005).

have been injured in the course of the stabbings.⁴⁸ But this man was not Stephen Watson—Watson’s DNA profile was different.⁴⁹

So the government dismissed the murder charge against Watson.⁵⁰ It also asked the state of Virginia to run the profile of the unknown samples through the state’s database of profiles from 101,905 convicted offenders.⁵¹ At that time, the database was limited to eight “STR loci”⁵²—locations in the genome where the lengths of a class of sequences vary from one person to another.⁵³ The state’s computer whizzed through the digits representing the length variations, or “alleles.” Out popped the name “Robert P. Garrett,” a known alias of Raymond Anthony Jenkins.⁵⁴

Jenkins was easy to find. Normally, he lived across the Potomac River, in Arlington, Virginia, but having been arrested on an unrelated burglary charge a few weeks after Dolinger’s death, he was in prison.⁵⁵ The authorities secured a search warrant for his blood.⁵⁶ To no one’s surprise, it matched the blood from Dolinger’s house. Indeed, the fresh sample provided a full thirteen-locus match.⁵⁷ The government’s estimates of the frequencies of the profile in the general population were astonishingly small—between one in twenty-six quintillion (1/26,000,000,000,000,000,000) and one in four sextillion (1/4,000,000,000,000,000,000,000).⁵⁸ In January 2000, Jenkins was charged with the murder.⁵⁹

48. *Jenkins*, 887 A.2d at 1017.

49. *Id.*

50. *Id.* at 1017 n.3.

51. *Id.* at 1017.

52. *Id.* at 1017 n.4.

53. Although an eight-locus match typically is quite revealing, *see, e.g.*, NAT’L COMM’N ON THE FUTURE OF DNA EVIDENCE, THE FUTURE OF FORENSIC DNA TESTING: PREDICTIONS OF THE RESEARCH AND DEVELOPMENT WORKING GROUP (2000), *available at* <http://www.ncjrs.gov/pdffiles1/nij/183697.pdf>, the standard number of loci used in databases in the United States currently is thirteen. *See, e.g.*, FBI, CODIS: COMBINED DNA INDEX SYSTEM (brochure), *available at* http://www.fbi.gov/hq/lab/html/codisbrochure_text.htm. Very roughly, each additional locus increases the discriminating power of a match by a factor of ten. *See* NAT’L COMM’N ON THE FUTURE OF DNA EVIDENCE, *supra*.

54. *Jenkins*, 887 A.2d at 1017.

55. Henri E. Cauvin, *Jury Deadlocks in Case that Relied on DNA*, WASH. POST, Apr. 15, 2005, at B10.

56. *Jenkins*, 887 A.2d at 1017.

57. *Id.*

58. *Jenkins*, 887 A.2d at 1019. Stated more succinctly, the estimated random-match probability in various racial or ethnic groups was on the order of 10^{-19} to 10^{-21} .

59. Weiss, *supra* note 38.

1. A Battle over the DNA Match

Jenkins filed a motion in limine to exclude the DNA evidence.⁶⁰ Of the various attacks on its admissibility, one struck home. Jenkins argued that presenting the numbers quoted above to express the significance of a DNA match in a trawl case is not generally accepted in the scientific community and therefore is inadmissible under *Frye v. United States*. There is, he contended, a “raging debate” in the scientific and statistical community regarding the most appropriate method for calculating the significance of a cold hit.⁶¹ He had a point. In the words of one statistician, “[r]arely can there have been such an important application of statistics in which . . . differing intuitions and approaches . . . lead to answers so vividly and violently opposed.”⁶²

The government argued that the debate was beside the point. The disagreement, it maintained, concerned the relevance of various probabilities or statistics rather than the validity of the formulas for computing them.⁶³ And relevance, it insisted, was a matter for the court, not the scientific community, to decide.⁶⁴

The trial judge held a pretrial hearing on the objection. Besieged with, “among other things, scholarly articles from reputable professional journals[, . . expert affidavits from leaders in the fields of genetics and statistics[, and] live testimony of” a human geneticist and two population geneticists,⁶⁵ Judge Rhonda Reid Winston struggled valiantly to describe and categorize the cacophony of expert voices.⁶⁶ She found that no method or combination of methods for presenting probabilities in a trawl case was generally accepted. It was “the last thing the Court wants, but I think that that’s where I am

60. *Jenkins*, 887 A.2d at 1017.

61. *Id.*

62. Philip Dawid, *Bayes's Theorem and Weighing Evidence by Juries*, in *BAYES'S THEOREM* 71, 87 (Richard Swinburne ed. 2002).

63. Transcript of Record, *supra* note 47, at 5.

64. *Id.*

65. *United States v. Jenkins*, 887 A.2d 1013, 1017–18 (D.C. 2005).

66. The prosecution and defense traded affidavits solicited from members of the 1996 committee giving their interpretations of the NRC report's recommendations. *See id.* at 1017 (noting that “[b]oth parties submitted . . . expert affidavits from leaders in the fields of genetics and statistics”). I submitted an affidavit for the defendant stating that I did not consider it appropriate to attempt to divine, years later, what the various committee members thought the words of the report meant. I pointed out, however, that as evidenced by the responses in the legal and statistical literature, the readers of the report at the time generally understood it as calling for the presentation of an adjusted probability when the defendant was identified through a database search. More recently, yet another member of the committee told newspaper reporters that “[t]he intent was to replace Random Match Probability.” Felch & Dolan, *supra* note 4.

forced to end up.”⁶⁷ Rejecting the government’s position that *Frye* did not apply and assuming that some probability had to be presented, she threw out all the DNA evidence.⁶⁸

In an interlocutory appeal, the District of Columbia Court of Appeals overruled the order.⁶⁹ It adopted the government’s view that *Frye* need not be applied.⁷⁰ The case went to trial with the random-match probability.⁷¹ But even this evidence was not enough. After five days of deliberations, the jury remained deadlocked, and the superior court declared a mistrial.⁷² The government elected to retry Jenkins. This time, the jury convicted Jenkins of all charges, including felony first-degree murder. “We’re just so elated,” said Michele Gehrke, a cousin of Dolinger’s.⁷³

2. The Easy Way Out: The p_{rem} and np Rules

In trawl cases like *Jenkins*, the admissibility issues are twofold. First, is a particular quantity in the case of a database search a generally accepted statistic or a scientifically valid one? Four quantitative measures of significance for a match coming out of a database trawl might be proposed: the frequency (or “rarity”) of a DNA profile in the relevant population, the random-match probability (which equals the expected frequency), the “likelihood ratio” (which is given by the reciprocal of the random-match probability when it is certain that the laboratory will declare a match between samples with the same profile),⁷⁴ and the “posterior probability” that the defendant is the source of the crime-scene sample.⁷⁵ If the method for computing any of these quantities is not generally accepted, then the corresponding statistic must be excluded in *Frye* jurisdictions.⁷⁶ If the method is not scientifically valid, then the statistic it produces must be excluded in *Daubert* jurisdictions.⁷⁷ Even if the *Frye* or *Daubert* hurdle is cleared, however, courts must also determine whether the statistic is irrelevant or unduly

67. Transcript of Record, *supra* note 47, at 25.

68. *Jenkins*, 887 A.2d at 1020.

69. *Id.* at 1026.

70. *Id.* (“We hold only that the trial court erred in subjecting a debate on relevancy to the exacting *Frye* standard of general acceptance in the relevant scientific community.”).

71. Cauvin, *supra* note 55.

72. *Id.*

73. Weiss, *supra* note 38.

74. The technical concept of “likelihood” is explained *infra* Parts II & III.

75. “Posterior” and “prior” probabilities are explained *infra* Parts II & III.

76. See KAYE ET AL., *supra* note 18, § 5.3.3.

77. *Id.*

misleading. If so, the quantity (and perhaps the underlying match) must be excluded under the familiar rules that require all evidence to be relevant and not unfairly prejudicial.⁷⁸

Initially, the government's effort to introduce the DNA match in *Jenkins* foundered when the trial court perceived no way to avoid the conflicting scientific literature on the need to adjust the random-match probability in trawl cases.⁷⁹ Among other things, both NRC reports recommend adjusting the random-match probability when the defendant is selected for prosecution as a result of the trawl. The 1992 NAS committee explained that

[t]he distinction between finding a match between an evidence sample and a suspect sample and finding a match between an evidence sample and one of many entries in a DNA profile databank is important. The chance of finding a match in the second case is considerably higher, because one does not start with a single hypothesis to test (i.e., that the evidence was left by a particular suspect), but instead fishes through the databank, trying out many hypotheses.⁸⁰

The committee observed that "there are statistical methods for correcting for such multiple testing," but proposed "a far better solution" to this problem: "When a match is obtained between an evidence sample and a databank entry, the match should be confirmed by testing with additional loci. . . . [O]nly the statistical frequency associated with the additional loci should be presented at trial (to prevent the selection bias that is inherent in searching a databank)."⁸¹

This procedure may be designated the "limit the loci" approach because it reserves loci for further testing after an initial set of hits in the database and ignores the match at the initial set of loci in computing the match probability. The adjusted probability, p_{rem} , which is based solely on the remaining loci, will be larger—probably orders of magnitude larger—than the random-match probability p for all the loci. This larger probability implies that the match is less probative. To appreciate this, suppose that a single match turns up when one searches the database at a predetermined set of eight loci

78. See, e.g., FED. R. EVID. 401, 403. In general, this Article does not address the issue of the prejudicial impact of a generally accepted or scientifically valid figure. For cases and commentary on this point, see, for example, *State v. Bloom*, 516 N.W.2d 159 (Minn. 1994), and KAYE ET AL., *supra* note 28, § 12.

79. *Jenkins*, 887 A.2d at 1020.

80. NRC I, *supra* note 15, at 124.

81. *Id.*

and that the suspect so identified matches at the remaining five loci used to confirm the preliminary match. If the random-match probability for all thirteen loci is one in ten trillion (10^{-13}) and the random-match probability for the five confirmatory loci is one in one hundred thousand (10^{-5}), the jury would receive only the figure of one in one hundred thousand for the probability of a matching profile in the general population.

Dismissing the bulk of the match (the eight loci used in the trawl) is generally a very conservative way to “correct” for the “multiple testing” that worried the 1992 committee. From the frequentist perspective, however, it is appropriate because it indicates how often database trawls could lead to false inclusions. If the two-stage procedure recommended in the 1992 NRC report were followed in a vast number of cases involving DNA profiles that give rise to a five-locus random-match probability of one in one hundred thousand, and if no database ever contains DNA from the true source of the crime-scene DNA, then the rate at which suspects would be falsely identified from their DNA would not exceed one in one hundred thousand.⁸²

The 1996 NAS committee started from the same premise that giving an upper bound on the rate of false identifications is important. “There is,” the committee maintained, “an important difference [when] the suspect is initially identified by searching a database to find a DNA profile matching that left at a crime scene.”⁸³ The committee then wrote that “the calculation of a match probability . . . should take into account the search process” and that “there are different ways to take the search process into account.”⁸⁴ Its preferred “correction to account for the database search” was to “multiply the match probability [p] by the size [n] of the database searched.”⁸⁵ With a trawl through a national database as large as the ones in the United Kingdom or the United States, this multiply-by-the-size-of-the-database approach would mean multiplying by a factor of about six million. A random-match probability of one in ten

82. One in one hundred thousand is an expected value that applies in the limit as the number of trawls approaches infinity. For a finite numbers of trawls, the error rate could be somewhat larger or smaller, just as the proportion of heads in a large but finite set of tosses of a fair coin need not be exactly one-half. See, e.g., MORRIS H. DEGROOT & MARK J. SCHERVISH, *PROBABILITY AND STATISTICS*, 233–35 (3d ed. 2002) (proving the law of large numbers that underlies the claim that the mean value of a variable in a large random sample from an infinite population is likely to be close to the population mean).

83. NRC II, *supra* note 16, at 134.

84. *Id.*

85. *Id.* at 134–35.

trillion ($1/10,000,000,000$) would increase to about six in ten thousand ($6/10,000$). The latter number is an indication of how often searching “innocent databases” (those that never contain the true source) of size $n = 6,000,000$ for DNA profiles with random-match probabilities of $p = 1/10,000,000,000$ will generate cold hits that could result in charges against an innocent person.

Several astute statisticians promptly challenged the committees’ recommendation to adjust the usual random-match probability p . Their views are described more fully in Part II.B. Given the very vocal disagreement of respected scientists and statisticians, the trial judge in *Jenkins* accurately perceived a serious controversy.

Nonetheless, there was an easy way out of the controversy. Judge Winston could have allowed the government to present an adjusted number that the scientific community properly regards as either fair or very generous to the defendant. Both the limit-the-loci approach and the np rule fill this bill. Unintentionally, the database search in *Jenkins* proceeded with a limited number of loci, as contemplated by the 1992 NAS committee. The initial database match was limited to eight loci, and another five were added only after the verification sample of Jenkins’ blood was obtained with a search warrant.⁸⁶ Rather than present the full thirteen-locus random-match probability, the government could have provided the random-match probability for the five-locus STR type that played no role in the database trawl. This value of p_{rem} would have been far larger than the government’s estimate of quadrillionths and sextillionths for the thirteen-locus random-match probability. Indeed, the probability as adjusted to the limited set of five loci could have been as large as $1/1634$.⁸⁷ With match probabilities of this magnitude, a defendant is in a much better position to maintain that DNA evidence leaves room for reasonable doubt.

The trial judge was unwilling to countenance the introduction of p_{rem} , however, because limiting the match probability to these loci (1) “wasn’t ever implemented [before],” (2) “wastes evidence,” and because (3) “everybody is testing the maximum number of loci that

86. This assumes that the thirteen-locus match of the second sample from Jenkins included all of the eight loci from the original match.

87. The figure of $1/1634$ is the largest random-match probability for the most common alleles at any five loci in any of the three groups (Caucasians, African Americans, and Hispanics) studied in John M. Butler et al., *Allele Frequencies for 15 Autosomal STR Loci on U.S. Caucasian, African American, and Hispanic Populations*, 48 J. FORENSIC SCI. 908, 909–11 (2003).

are available [s]o it would be difficult to implement.”⁸⁸ However, none of these explanations constitute a reason not to have allowed the five-locus probability. It is true that laboratories have not rushed to limit the number of loci used in trawling, but the software for trawling allows searches at a reduced number of loci. The software has to have this capacity because crime-scene samples can be too degraded or minute to be typeable at all thirteen core CODIS loci.⁸⁹ In addition, sample size permitting, it is possible to examine loci beyond the thirteen standard ones after an initial database match at all thirteen.⁹⁰ Therefore, contrary to the third point raised by the trial judge, the limit-the-loci approach would not be difficult to implement in general. Moreover, having unintentionally conducted the initial search in the Virginia database in the manner prescribed by the NAS committee, the adjusted number certainly was available in *Jenkins*.

Neither are points (1) or (2) persuasive. Despite the lack of implementation and the waste of information, there is no doubt that the limit-the-loci approach is a conservative response to the frequentist objection to random-match probability in a trawl case. As Judge Winston noted, “it’s conservative and nobody disagreed with the method that was used to calculate it.”⁹¹ Because there is no dispute that the limit-the-loci approach is a scientifically valid way to arrive at a probability that favors the defendant, a defendant cannot validly contend that *Frye* or *Daubert* bars the introduction of p_{rem} . Thus, Judge Winston could have disposed of the *Frye* objection by admitting the match along with only the adjusted figure p_{rem} .⁹²

88. Transcript of Record, *supra* note 47, at 25.

89. The “core loci” are those that the FBI designated for use in the National DNA Index System (NDIS). See, e.g., FBI’s DNA Advisory Board, *supra* note 22, at 94.

90. The real reason not to require second-stage additional tests is that they would be a waste of laboratory time and effort. The identifying information in a thirteen-locus match is more than sufficient even in a trawl case. See *infra* Part II.

91. Transcript of Record, *supra* note 47, at 24–25; cf. N. E. Morton, *The Forensic DNA Endgame*, 37 JURIMETRICS J. 477, 489 (1997) (“To the delight of scientists and judges and the disappointment of mathematicians, this solution puts an end to controversy about interpretation of suspect trawls in a very large database.”).

92. That said, it may be unfair to fault the trial court for not finding general acceptance when both parties urged this view of the limit-the-loci approach on the court. Assistant U.S. Attorney Michael Ambrosino advised the court that:

[O]ur position, Your Honor, is that I don’t believe there is general acceptance. You can[not] find general acceptance of a principle that’s never been . . . implemented by any forensic lab. . . . [N]o, I don’t think the Court could hold there is general acceptance of something that’s never been done even since it’s [sic] recommendation in ‘92.

Transcript of Record, *supra* note 47, at 24.

Alternatively, the court could have admitted the *np* figure. As with the limit-the-loci approach, the complaint about the *np* rule in the forensic statistical literature is that it is *unnecessarily* conservative.⁹³ This may make the *np* statistic unappetizing to the prosecution, which would prefer the smaller random-match probability, but it is not a reason to exclude it under *Frye* or *Daubert*. Once again, a defendant cannot coherently argue that a probability estimate that scientists agree is either correct or too large (and hence understates the power of the DNA evidence) cannot be introduced against him.

In *Jenkins*, application of the *np* rule actually would have made no meaningful difference in the match probability. The unadjusted probability was, at most, about $p = 10^{-19}$, and the database size was about $n = 10^5$. The *np* rule therefore gives a match probability of approximately 10^{-14} , or one in a hundred trillion. Nonetheless, the government was reluctant to give this figure to the jury even as a supplement to the unadjusted value of *p*. The ostensible reason was solicitude for the rights of the defendant. At oral argument, “[t]he government indicated that . . . it would [not] present both . . . statistics . . . without a trial court request [because of] the potential prejudicial effect of informing the jury that Mr. Jenkins’ DNA profile was already on file in a Virginia offender database.”⁹⁴

The court of appeals was unimpressed with this explanation. It responded that

[t]his prejudice, however, is no different than that which would have occurred had the trial court allowed introduction of statistics derived from the 1992 National Research Council recommendation to test confirmatory loci. In explaining why “untainted” loci were tested, the proffered expert would be required to state that Mr. Jenkins was first identified through a search of an offender database. Because countless trial courts, including the trial court in the instant case, have so readily accepted the statistics derived from testing untainted loci, we must surmise that conveying to the jury that the defendant was

Still, it is unfortunate that the court did not recognize that the position of all counsel in the case was untenable. The reluctance of the DNA-testing community to implement the procedure does not reflect any disagreement over the proposition that it produces a generous quantitative measure of the probative value of a match in a trawl case. The main reason that the limit-the-loci approach has not been implemented is that prosecutors prefer the smallest probabilities they can get and laboratories have no incentive to adopt a procedure that they see as understating the value of their findings.

93. See, e.g., Balding, *supra* note 11, at 470–73.

94. United States v. Jenkins, 887 A.2d 1013, 1019 n.13 (D.C. 2005).

first identified through a search of an offender database has not been deemed so substantially prejudicial as to outweigh the probative value of such evidence.⁹⁵

This categorical rejection of the character evidence concern is unconvincing because it assumes that adjusting the random-match probability necessarily entails revealing that the defendant has a prior record. In fact, the expert can testify that the probability of a match if the crime-scene DNA came from a person genetically unrelated to the defendant is the adjusted number (p_{rem} or np) or less. There is no need to explain on direct examination that the “or less” qualification stems from the trawl, and it is not clear that revelation of trawling has occurred in “countless” cases.

At the same time, a categorical rule against revealing that a defendant’s DNA was in a law enforcement database also seems too strong. The general rule against introducing evidence of a defendant’s other crimes does not apply when the purpose of the revelation is something other than the suggestion of a general propensity to criminality.⁹⁶ Accordingly, this rule should not prevent the prosecution (and certainly not the defense) from revealing that the defendant was found through a DNA database trawl, at least where the bona fide purpose of the information is to help the jury to understand the DNA match. Instead of a categorical rule of admission or exclusion, a case-by-case balancing of the value of the information for this legitimate purpose as against its potential prejudice to the defendant is appropriate.⁹⁷

In short, the objection to the np figure in *Jenkins* could have been disposed of without prejudicing the defendant by presenting only that number (as an upper bound on the probability of a random match with anyone unrelated to the defendant). Why was the prosecution unwilling to present np in lieu of its more cherished values of quintillions and sextillionths? Surely, no prosecutor could have been distressed at the prospect of going into court with a “one in hundred trillion” match statistic.

One possibility is that the government was concerned about other cases involving highly degraded or limited DNA samples. With only a few alleles, np would be less overwhelming. In *People v. Puckett*, for instance, the random-match probability for the match at

95. *Id.* at 1025 n.20.

96. *See, e.g.*, 1 MCCORMICK ON EVIDENCE § 190 (Kenneth Broun ed., 6th ed. 2006).

97. *Id.*

six loci⁹⁸ was said to be about one in 1.1 million.⁹⁹ The California database had 338,000 profiles in it, making np almost one in three—a number that would render the match almost worthless to the prosecution.¹⁰⁰ If *Jenkins* was a test case for the supremacy of the unadjusted probability p , then the government succeeded brilliantly. The court of appeal opinion in *Jenkins* states that p and np are each somehow relevant in a trawl case and that neither is prejudicial.¹⁰¹

3. I've Got Your Numbers

Section A.2 showed that in *Jenkins* and in most trawl cases, neither *Frye* nor *Daubert* is a real barrier to admitting one or the other adjusted-match probabilities. Both adjustments favor the defendant and one is not inherently superior to the other. If the prosecution's experts are content to use only the np or p_{rem} statistics, a court need not try to judge the statistical debate over which figure logically applies. At worst, these numbers err on the side of the defendant. Because this evidentiary compromise was not acceptable to the parties in *Jenkins*, however, the District of Columbia Court of Appeals had to address the statistical controversy over adjustment of the random-match probability. This Subsection first explains how the court of appeals correctly concluded that the controversy over which statistic to use should not be analyzed in terms of the general-acceptance (or scientific-validity) standards. Having properly sited the issue in the general doctrine of relevance, however, the *Jenkins* court skirted the statistical debate, reasoning that all three statistics (p , np , and p_{rem}) are admissible because each one is relevant to some proposition.¹⁰² This Subsection argues that this attempt to sidestep the statistical debate is a move in the wrong direction; to resolve the

98. Felch & Dolan, *supra* note 4, refer to “5 ½” matching “genetic markers” that were “mixed with traces of DNA from another person, probably [the victim].” *Id.* The match probability introduced at trial came from six loci. Human beings have one or two STR types at each locus, depending on whether they inherited the same allele from each parent (homozygosity) or a different allele from each parent (heterozygosity). The laboratory reported clear STR types at four loci (one being homozygous). A fifth locus showed three alleles, indicating a mixture consistent with Puckett and the victim. A sixth locus showed two alleles, but one gave too weak a signal to be characterized. S.F. POLICE DEP'T, FORENSIC SCIENCE DIVISION, CRIMINALISTICS LABORATORY, SUPPLEMENTAL REPORT OF LABORATORY EXAMINATION, CASE NO. 040 027 547 & 721 032 24(0), Aug. 24, 2007, at 3 (on file with the North Carolina Law Review).

99. S. F. POLICE DEP'T, *supra* note 98, at 4; Felch & Dolan, *supra* note 4.

100. These figures are taken from Felch & Dolan, *supra* note 4.

101. *United States v. Jenkins*, 887 A.2d 1013, 1024–25 (D.C. 2005).

102. *Id.*

relevance question, courts must confront the disagreement among statisticians.

To begin with, the court of appeal in *Jenkins* correctly identified the logical nature of the statistical dispute:

At the heart of this debate is a disagreement over the competing questions to be asked, not the methodologies used to answer those questions. . . . T]here is no controversy in the relevant scientific community as to the accuracy of the various formulas. In other words, the math that underlies the calculations is not being questioned. Each approach to expressing significance of a cold hit DNA match accurately answers the question it seeks to address. The rarity statistic [p] accurately expresses how rare a genetic profile is in a given society. Database match probability [np] accurately expresses the probability of obtaining a cold hit from a search of a particular database. [Another computation] accurately expresses the [posterior] probability that the person identified through the cold hit is the actual source of the DNA in light of the fact that a known quantity of potential suspects was eliminated through the database search. These competing schools of thought do not question or challenge the validity of the computations and mathematics relied upon by the others. Instead, the arguments raised by each of the proponents simply state that their formulation is more probative, not more correct. Thus, the debate cited by Mr. Jenkins is one of relevancy, not methodology; and . . . there is no basis under *Porter* for the trial court to exclude the DNA evidence in this case.¹⁰³

Concluding that “*Frye*, therefore, does not impede introduction of the statistics into evidence,”¹⁰⁴ the *Jenkins* court moved on to the real question in the case—which number or numbers are relevant after a trawl? Despite this promising start, the court’s answer to the relevance question is disappointing. The court insisted that

the most probative evidence for a factfinder is that which indicates whether or not the DNA obtained from a crime scene matches the DNA obtained from the suspect. The likelihood that the suspect is the actual source of the DNA is best expressed through the rarity of a particular profile. Thus, the

103. *Id.* at 1022–23 (citations omitted). “*Porter*” refers to *United States v. Porter*, 618 A.2d 629 (D.C. 1992). In that case, the District of Columbia Court of Appeals held that a particular procedure for computing a random-match probability in a confirmation case did not satisfy *Frye* and directed the trial court to consider whether an alternative method was generally accepted as conservative. *Id.* at 631.

104. *Jenkins*, 887 A.2d. at 1026.

rarity statistic is highly probative and will always be relevant. In fact, in this appellate record there is nothing that would suggest that the probative value of . . . the rarity statistic . . . is substantially outweighed by its prejudicial impact.¹⁰⁵

The “rarity statistic” is simply a relative frequency. In *Jenkins*, it was one in so many quintillions or sextillions. In *Puckett*, it was one in 1.1 million. But why is “the rarity of a particular profile” the *best* measure of the probability “that the suspect is the actual source of the DNA”? Indeed, why is it even relevant to this proposition? The frequentist position is that the frequency, or rarity, is relevant *in a confirmation case* because the frequency is also the random-match probability, which can be understood as the probability of a matching profile when the defendant is not the source (and is not closely related to the actual source). In other words, the frequency is a *p*-value, and small *p*-values justify rejection of the null hypothesis that the suspect is not the actual source.¹⁰⁶ Once this null hypothesis is rejected, the jury is a step closer to the conclusion that the defendant is in fact the actual source.

But, defense counsel can argue this logic applies only to the confirmation case. In that kind of a case, the frequency, or random-match probability, is indeed the probability of the match to an innocent, unrelated person. With the frequency estimate of one in 1.1 million in *Puckett*, it would be quite surprising if Puckett were the only person tested and he matched. But Puckett was not the only person tested. Over three hundred thousand DNA profiles were examined. If everyone in the database were innocent, it still would be no great surprise to learn that one or more of them matched. The *p*-value for the trawl (as computed with the *np* rule) is about one in three. Hence, the frequentist perspective seems to suggest that the “rarity statistic” is *not* the best measure of probative value. On the contrary, it is thoroughly misleading.

This seems like a powerful argument against the relevance of the “rarity statistic” and its equivalent, the random-match probability, but the *Jenkins* court basically ignores it. Rather than join or assess the debate between the frequentists and their critics, the court conveniently assumes that, in a trawl case, the jury would benefit from the answers to three distinct questions: (1) “[H]ow rare [is] a genetic profile . . . in a given society” (the random-match probability)? (2) How probable is “a cold hit from a search of a

105. *Id.* at 1025 (citations omitted).

106. *See supra* note 19 and accompanying text.

particular database” (the *np* statistic)? And (3) How probable is it “that the person identified through the cold hit is the actual source of the DNA in light of the fact that [the defendant’s DNA profile matched while] a known quantity of potential suspects was eliminated through the database search” (the posterior probability)?¹⁰⁷

But why are these questions relevant, and why is any one of them more apposite than another? Judicial fiat cannot supply the answer. Neither can the disciplines of population genetics, human genetics, or molecular biology. The issue is one of probability, statistics, and inductive logic. We need to analyze the probative value of a DNA match—the extent to which it supports the prosecution’s claim that the defendant is the source of the crime-scene DNA¹⁰⁸—with and without a trawl. That is the subject of both the articles and the debate that the *Jenkins* court deemed irrelevant. Without addressing the debate, the *Jenkins* court has no basis to conclude that the frequency or random-match probability *p* is preferable to or even admissible along with the database-match probability *np*. We shall analyze this debate, and defend the result in *Jenkins*, in Part II.

B. California Dreaming: Johnson and Nelson

The next challenges to random-match probabilities in database trawl cases that reached the appellate courts occurred in California. The California courts have followed *Jenkins*, but they also have devised a new theory as to why an unadjusted random-match probability can be admitted in a trawl case. This rationale can be called a “replication theory” since it holds that the problem with the trawl vanishes once the match is replicated with a fresh sample from the suspect whose name came from the trawl. This Section shows that the replication theory is patently fallacious.

The replication theory first took root in *People v. Johnson*. On a winter’s night in 1996, a fifteen-year-old girl was at a pay telephone in Visalia, California, speaking to her boyfriend.¹⁰⁹ A man speaking a mixture of Spanish and English drew a knife, forced her into a pickup truck, and drove to a rural area.¹¹⁰ He raped her, and then drove her back into town.¹¹¹ At a hospital, semen was collected from her body,

107. *Jenkins*, 887 A.2d at 1023.

108. On probative value generally, see 1 MCCORMICK, *supra* note 96, § 185, at 306–10. For a quantitative analysis, see D.H. Kaye & Jonathan J. Koehler, *The Misquantification of Probative Value*, 27 LAW & HUM. BEHAV. 645, 646–47 (2003).

109. *People v. Johnson*, 43 Cal. Rptr. 3d 587, 590 (Ct. App. 2006).

110. *Id.*

111. *Id.*

and she described the assailant and the truck in impressive detail.¹¹² After some initial efforts, the police gave up on the case.¹¹³

Five years later, the unanalyzed rape kit from the hospital examination went to the California Department of Justice Regional Laboratory.¹¹⁴ A criminalist detected sperm cells on one of the vaginal smear slides.¹¹⁵ She sent the materials to a central laboratory, which obtained a thirteen-locus STR profile.¹¹⁶ A trawl of the national database implicated a prisoner named Michael Johnson.¹¹⁷ It was now 2003.¹¹⁸ Investigators took a blood sample from Johnson.¹¹⁹ The DNA matched that in the sperm.¹²⁰ A criminalist estimated that the matching STR genotype would “occur at random in the general population in about one in 130 quadrillion African-Americans, one in 240 quadrillion Caucasians, and one in 4.3 quadrillion Hispanics.”¹²¹

After hearing this testimony, a jury convicted Johnson of rape and related offenses.¹²² He appealed. One of his arguments was that, at the time, STR typing and the random-match probability for an STR genotype were not generally accepted in the scientific community.¹²³ The trial court rejected this claim after a pretrial evidentiary hearing.¹²⁴ Although Johnson raised no particular objection in these proceedings to using a random-match probability in a trawl case, on appeal he pressed the point that “because the instant case involved a cold hit, the foundation concerning the statistical

112. *Id.* at 590–91.

113. The day after the assault, the police showed her a pickup truck that they had impounded. She was positive that it was the one in which she lay crouched on the floor boards as a tattooed, Hispanic male with black hair and a “Fu manchu mustache” drove her back and forth. *Id.* at 591. However, the truck’s owner was Robert Calkins, a white man “with long hair and a fullbrown [sic] beard.” *Id.* at 591 n.3. What the police apparently did not realize (or failed to appreciate) was that Calkins had just bought the truck, and it seems that they did not investigate who else might have been driving it. *Id.* at 591. The police showed the girl 576 photographs of possible suspects. She was unable to identify anyone. By late 1996, they gave up on the case. *Id.*

114. *Id.* at 591.

115. *Id.*

116. *Id.*

117. *Id.*

118. *Id.*

119. *Id.*

120. *Id.*

121. *Id.*

122. *Id.* at 589.

123. *Id.* at 590.

124. *Id.* at 592–93.

interpretation of the DNA evidence was insufficient to satisfy the [general acceptance] standard for scientific evidence.”¹²⁵

The California Court of Appeal rejected the claim, but it stated that “the presentation of various statistical analyses” of the three questions listed in *Jenkins* “would raise significant relevancy and [prejudice] issues.”¹²⁶ The court chose not to follow *Jenkins*’ reasoning that the question of which statistic to apply (p , p_{rem} , or np) is simply a matter of logical relevance for the court to determine. Instead, the court sought to avoid the thorny statistical question on the curious ground that “the database search merely provides law enforcement with an investigative tool, not evidence of guilt.”¹²⁷ General acceptance does not apply to the method of computing a match probability for a database search, because “the use of database searches as a means of identifying potential suspects is not new or novel” and “no authority [applies the] requirements to a mere investigative technique.”¹²⁸

It is true that the initial match in the database will be replicated by drawing and analyzing a new sample from the individual involved. This is what happened in *Jenkins* and *Johnson*. Such replication is a red herring, however, because the challenge is not to the use of a convicted-offender DNA database as an investigatory tool. The objection is to the use of the random-match probability at trial to gauge the power of the later match when the defendant has not been selected for DNA testing “at random”—that is to say, on the basis of factors that are uncorrelated with his DNA profile. When the defendant is selected for a later test precisely because of his known DNA profile, the replication adds no new information about the hypothesis that the defendant is unrelated to the actual perpetrator and just happens to have the matching DNA profile.¹²⁹ Replication

125. *Id.* at 594. Ordinarily, an appellate court would not entertain an objection to evidence that had not been raised at trial. The court of appeal thought that “the [procedural] question here is a close one” and elected to “assume the issue is cognizable and decide it on the merits to forestall appellant’s alternative claim that trial counsel was ineffective if he failed adequately to preserve the issue.” *Id.* at 594 n.11.

126. *Id.* at 601.

127. *Id.* at 597.

128. *Id.* at 596–97. The court expressed “no opinion concerning a situation in which the fact of the preliminary, ‘cold hit’ match from the offender database is offered as evidence of guilt.” *Id.* at 598 n.17.

129. It adds no such information because the datum—a matching profile in the new sample—is just as probable when this hypothesis is true as when it is false. If the defendant is innocent and just happens to have the incriminating profile (and if the laboratory reports are accurate) then he will have that profile in one sample after another. Likewise, if the defendant is guilty and has the incriminating profile because he is indeed

helps eliminate the risk of a laboratory error in determining or reporting the STRs, but it has no further value in probing the possibility of a coincidental match.

Yet the *Johnson* opinion proclaims that the original match, rather than the replicated match, is logically irrelevant:

In our view, the means by which a particular person comes to be suspected of a crime—the reason law enforcement’s investigation focuses on him—is irrelevant to the issue to be decided at trial, i.e., that person’s guilt or innocence, except insofar as it provides independent evidence of guilt or innocence. For example, assume police are investigating a robbery. The victim identifies “Joey” as the perpetrator. The means by which “Joey” becomes the focus of the investigation—the eyewitness identification—is relevant because that identification is itself evidence of guilt. Suppose instead that a surveillance camera captures the robbery on tape. Police use facial recognition software to check the robber’s facial features against driver’s license photographs. When the computer indicates a match with “Joey,” officers obtain his name and address from DMV records, then go to his house and interview him. In the course of the interview, “Joey” confesses. Whether facial recognition software is discerning and accurate enough to select the perpetrator, or whether it declared a match involving many different people who resembled “Joey,” or how many driver’s license photographs were searched by the software, is immaterial: what matters is the subsequent confirmatory investigation.¹³⁰

The *Johnson* court’s analogy is flawed because it assumes that Joey is no more likely to confess if he is found by the face-recognition software than by some other means. But suppose that whenever a suspect is identified by the computer program, the police pressure him into confessing (because they believe in the technology) and that they apply no such pressure in other situations (where they are less certain of guilt). Then the confession is redundant, and how Joey came to the attention of the police makes quite a difference in the probative value of the confession.

A more apt analogy would be a case in which Joey is identified by the pattern-recognition algorithm applied to the drivers’ licenses, and the police then take a fresh photograph of Joey and process it

the source of the crime-scene DNA, he will have the same profile in one sample after another.

130. *Johnson*, 43 Cal. Rptr. 3d at 597 (citations omitted) (emphasis omitted).

with the same scanner and algorithm. Naturally, there are two matches. But what of it? If a trawl through the DMV records degrades probative value, the confirmation with the new photograph does not restore what was lost. It is as if the police got their suspect from the trawl, then ran the program a second time with just the defendant's DMV photo. Contrary to the suggestion in *Johnson*,¹³¹ the second, redundant effort cannot make the fact of the initial trawl "irrelevant to the issue to be decided at trial."¹³² In short, the replication theory does not avoid the trawl objection. The objection is that the search strategy degrades the probative value of *both* the initial match in the database *and* the subsequent replication match, requiring an adjustment to the random-match probability to provide a scientifically valid and fair measure of either item of evidence.

Despite this fact, the replication theory adopted in *Johnson* spread to another of California's judicial districts. In *People v. Nelson*, the victim, a college student, had been abducted at a parking lot in 1976, brutally "raped and drowned in mud."¹³³ Dennis Louis Nelson was a suspect, but the police never developed an adequate case against him.¹³⁴ Twenty-five years later, semen derived from material in storage yielded an STR profile.¹³⁵ A trawl through the state's database netted Nelson.¹³⁶ The laboratory reported that "this profile would occur at random among unrelated individuals in about one in nine hundred and fifty sextillion African Americans, one in one hundred and thirty septillion Caucasians, and one in nine hundred and thirty sextillion Hispanics."¹³⁷

131. See *supra* note 129 and accompanying text.

132. *Johnson*, 43 Cal. Rptr. 3d at 597. The replication in these pattern-recognition cases—photographs or DNA—is completely unlike the kind of replication that scientists demand before they are inclined to accept as real the finding of an association between a locus found to be associated with a common disease by trawling across many loci. David J. Hunter & Peter Kraft, *Drinking from the Fire Hose—Statistical Issues in Genomewide Association Studies*, 357 NEW ENG. J. MED. 436, 436 (2007). In that case, if the original finding were mere happenstance, it would be surprising to find a strong association with the same locus in an independent sample. In contrast, it is no surprise that the laboratory can find the same profile when it examines a second or even a third sample from the same individual. That kind of replication should occur whether or not the original finding was happenstance. With a likelihood ratio of one (with respect to the hypothesis of a coincidental match among unrelated individuals), it is the later evidence that is arguably irrelevant.

133. *People v. Nelson*, 185 P.3d 49, 53 (Cal. 2008).

134. *Id.*

135. *Id.*

136. *Id.*

137. *People v. Nelson* 48 Cal. Rptr. 3d 399, 404 n.2 (Ct. App. 2005).

In 2002, Nelson was charged with the 1976 rape and murder. After a pretrial effort to suppress the DNA evidence failed, he maintained at trial that he “had consensual intercourse [with the victim] on the weekend before she disappeared and that someone else abducted, raped, and murdered her.”¹³⁸ The defense was “without evidentiary support,”¹³⁹ and the jury convicted Nelson of murder.

Like Johnson and Jenkins before him, Nelson argued on appeal that “the DNA evidence should not have been introduced because . . . there is no generally accepted statistical method for explaining the significance of DNA evidence when a suspect is identified through use of a convicted offender databank.”¹⁴⁰ Weaving together the flawed strands of reasoning in *Johnson* and *Jenkins*, the California Court of Appeal rejected the argument. The *Nelson* court wrote that

[u]se of the product rule to compute a random-match probability is overwhelmingly accepted by the scientific community as a scientifically reliable means of demonstrating the rarity of a profile in the population. . . . Consequently, the technique satisfies the reliability prong of the [general-acceptance] test and it was for the trial court, not the scientific community, to determine the relevance of the technique to this criminal prosecution.¹⁴¹

Unfortunately, like the D.C. Court of Appeals in *Jenkins*, the California Court of Appeal in *Nelson* provided no explanation of why “rarity” is relevant or probative not merely in a confirmation case, but also in a trawl case.

The other strand of the *Nelson* court’s analysis came from *Johnson*. The court insisted that

[t]he DNA databank search merely identified defendant as a possible candidate as the murderer; it was not the basis for declaring that his DNA matched DNA on the evidentiary samples. The latter determination was made based upon further, complete testing utilizing scientific techniques found to be reliable and admissible under the [general-acceptance] test.¹⁴²

The “further, complete testing,” however, only consisted of verifying the hit in the database with DNA taken from the defendant

138. *Id.* at 404.

139. *Id.*

140. *Id.* at 411.

141. *Id.* at 418.

142. *Id.* at 402.

with a warrant procured as a direct result of that cold hit. As discussed in connection with *Johnson*,¹⁴³ the argument that later searches replicate the match from the trawl is not responsive to the concern that an initial trawl dilutes the probative value of the matching DNA. The frequentist argument, it will be recalled, is that any number of DNA matches based on a trawl is weaker evidence than one DNA match to a suspect identified independently of his DNA type. The replication here is too transparent a garment to cloak the fact of a trawl.

On appeal, the California Supreme Court unanimously affirmed the Court of Appeal. The supreme court's opinion studiously (and perhaps pointedly) avoids the replication theory. Quoting from *Jenkins*, the California Supreme Court correctly reasoned that *Frye* does not apply to the question of which statistic is logically relevant.¹⁴⁴ Then the court relied on the fact that there was no dispute over the random-match probability as an estimate of "the rarity statistic."¹⁴⁵ However, the analysis remains incomplete, for the opinion does not discuss why "the rarity statistic" is relevant.¹⁴⁶ This population frequency or random-match probability is not relevant for its own sake, even in a simple confirmation case. It is relevant only insofar as

143. See *supra* notes 130–32 and accompanying text.

144. *People v. Nelson*, 185 P.3d 49, 61 (Cal. 2008) ("[I]t is already established that the product rule reliably shows what it purports to show—the rarity of the genetic profile in the population group. Accordingly, its admissibility in a cold hit case is a question of relevance, not scientific acceptance, and it is thus not subject to a further *Kelly* [California's version of *Frye*] test.").

145. *Id.* at 64.

146. *Nelson* goes one step beyond *Jenkins* in this regard but then stumbles. Justice Chin's opinion for the California Supreme Court states that:

The database match probability ascertains the probability of a match from a given database [that does not contain a profile from the source of the crime-scene sample]. 'But the database is not on trial. Only the defendant is.' ([D.H. Kaye & George Sensabaugh, *DNA Typing*, in 4] Modern Scientific Evidence[: The Law and Science of Expert Testimony] § 32:11, pp. 118–119 [David L. Faigman et al. eds., 2006].) Thus, the question of how probable it is that the defendant, not the database, is the source of the crime scene DNA remains relevant. (*Id.* at p. 119.) The rarity statistic addresses this question.

Id. at 66. In the treatise being quoted, however, these remarks describe an argument, developed more fully below, for using the *unadjusted* random-match probability *instead of* the adjusted probability. If "the database is not on trial," then the probability that an innocent database would be incriminated by the trawl is irrelevant and confusing. Yet, the *Nelson* opinion asserts that the adjusted statistic might be admissible along with the normal random-match probability. See *id.* at 66 n.3. This suggestion is discussed *infra* Part III.B.

it conveys the meaning of the match in the applicable context.¹⁴⁷ Whether the random-match probability, or any other statistic, is admissible to explain the meaning of one or more matches *resulting from trawling a database* depends on how well the statistic serves as an indicator of the probative value of the match that comes from the trawl. That the frequency or random-match probability is suitable in a confirmation case is not enough to justify its use in a trawl case. As the NRC committees and other commentators have contended, there is a logical difference between a confirmation and a trawl match.¹⁴⁸ That logical difference must be considered before one can conclude that the random-match probability is a valid and fair measure of probative value in trawl cases. We will have to venture into the statistical thicket after all. Only then can we discern whether an adjustment to the random-match probability is required in the case of a trawl, and if so, what it should be.

II. TRAWLING IN DATABASES: PROBATIVE VALUE

By avoiding the merits of the statistical debate over the effect of a database trawl, the opinions in *Jenkins*, *Johnson*, and *Nelson* are vulnerable to the criticism, developed in Part I, that they allow the use of a statistic that is helpful in the confirmation setting but irrelevant or misleading in the trawl situation. This Part rebuts that criticism. It directly confronts the claim that a trawl necessitates an adjustment to the random-match probability if the statistic is to be a valid indicator of probative value. Recognizing that there is a logically important distinction between trawl cases and confirmation cases, the analysis builds on a fundamental result in probability theory, Bayes' rule, to demonstrate that the frequency or random-match probability is as useful in trawl cases as it is in confirmation ones. To reach this conclusion, Section A provides a more complete explanation of the frequentist skepticism about the probative value of

147. See D. H. Kaye, *The Role of Race in DNA Evidence: What Experts Say, What California Courts Allow*, 37 S.W. L. REV. 303, 311–13 (2008) (explaining how jurors can use DNA statistics).

148. See, e.g., DAVID J. BALDING, *WEIGHT-OF-EVIDENCE FOR FORENSIC DNA PROFILES* 36 (2005) (contending that “DNA evidence is usually slightly stronger in the database search setting than when no search has occurred”); NRC I, *supra* note 15, at 124 (“The distinction between finding a match between an evidence sample and a suspect sample and finding a match between an evidence sample and one of many entries in a DNA profile databank is important.”); NRC II, *supra* note 16, at 134 (“There is an important difference [when] the suspect is initially identified by searching a database to find a DNA profile matching that left at a crime scene.”); *supra* note 80 and accompanying text.

a match that comes from a database trawl. Section B outlines the version of Bayes' rule that shows why the frequentist concern is misplaced in this context. Sections C and D use the mathematical formalism to compare the impact of trawling in various situations to a single, confirmation match. Section E defends the Bayesian analysis against some possible objections.

A. *The Dangers of Multiple Hypothesis Testing*

Statisticians are taught to be wary of "multiple hypothesis testing," "searching for significance," "ascertainment bias," or "data mining."¹⁴⁹ The process of subjecting drugs to clinical trials before accepting them as effective illustrates the problem. The best clinical trial is a randomized, controlled, double-blind experiment.¹⁵⁰ Patients are randomly divided into control and treatment groups.¹⁵¹ The control patients are given a harmless placebo, while the treated patients receive the experimental drug.¹⁵² To avoid any bias, during the trial period neither the subjects nor the researchers know which patients are in which group.¹⁵³ If there is a significant difference in the outcomes experienced by the treated patients as compared to the controls, then the statistical hypothesis test rejects the "null hypothesis," and we can conclude that the drug is effective.¹⁵⁴

149. *E.g.*, DAVID S. MOORE & GEORGE P. MCCABE, *INTRODUCTION TO THE PRACTICE OF STATISTICS* 427 (5th ed. 2006).

150. *See, e.g.*, HANS ZEISEL & DAVID KAYE, *PROVE IT WITH FIGURES: EMPIRICAL METHODS IN LAW AND LITIGATION* 2–3 (1997); *cf.* 21 C.F.R. § 314.126 (2008) (describing "[a]dequate and well-controlled studies" for the "purpose of conducting clinical investigations of a drug" as required by the FDA for marketing approval and stating in § 314.126(e) that "[u]ncontrolled studies or partially controlled studies are not acceptable as the sole basis for the approval of claims of effectiveness"); DONALD A. BERRY, *STATISTICS: A BAYESIAN PERSPECTIVE* 68 (1996) ("Randomization has obvious virtues in medical research."); Stuart J. Pocock & Diane R. Elbourne, *Randomized Trials or Observational Tribulations?*, 342 *NEW ENGL. J. MED.* 1907, 1908 (2000).

151. *E.g.*, 21 C.F.R. § 314.126(b)(4); MOORE & MCCABE, *supra* note 149, at 200–03; Philip W. Lavori et al., *Designs for Experiments: Parallel Comparison of Treatment, in MEDICAL USES OF STATISTICS* 61, 61 (John C. Bailar III & Frederick Mosteller eds., 2d ed. 1992).

152. MOORE & MCCABE, *supra* note 149, at 200; Lavori et al., *supra* note 151, at 61. This is the simplest experimental design. Many variations are possible. For example, investigators can create additional treatment groups to test for dosage effects. *E.g.*, 21 C.F.R. § 314.126(b)(2)(ii).

153. MOORE & MCCABE, *supra* note 149, at 206; *cf.* 21 C.F.R. § 314.126(b)(5) (requiring "[a]dequate measures . . . to minimize bias on the part of the subjects, observers, and analysts of the data, . . . such as blinding"); ZEISEL & KAYE, *supra* note 150, at 109 (concluding that "the double-blind rule constitutes a requirement of good, controlled experimentation with human subjects").

154. MOORE & MCCABE, *supra* note 149, at 203, 424.

There is no single, agreed-upon threshold for how much of a difference is needed to accept the claim of efficacy.¹⁵⁵ Even if a drug has no therapeutic value, random assignment of patients to each group sometimes places in the treatment group more patients who show improvement for reasons unrelated to the drug. To avoid declaring too many “false positives” in clinical trials, statisticians assume that the drug accomplishes nothing and compute the probability of seeing a difference as large or larger than the one actually observed in a clinical trial given that the differences reflect only the variation due to the random assignment of patients to the two groups.¹⁵⁶ If we demand that this probability—a *p*-value—be less than 0.05, then, in the long run, we will falsely declare efficacious no more than one in twenty drugs that go through clinical trials.¹⁵⁷

Why 0.05? This particular level usually is taken to denote statistical significance because of the remarks and tables of the eminent British statistician, Sir Ronald Fisher.¹⁵⁸ But the choice is somewhat arbitrary. If an expected false-positive error frequency of one in twenty seems too high, we can adopt a lower (more

155. *E.g., id.* at 424.

156. James H. Ware et al., *P Values*, in *MEDICAL USES OF STATISTICS*, 181, 183 (John C. Bailar III & Frederick Mosteller eds., 2.d ed. 1992).

157. *See, e.g.,* MOORE & MCCABE, *supra* note 149, at 438; Ware, *supra* note 156, at 189 (describing the 0.05 level).

158. Sir Ronald Fisher, a statistician and geneticist at the agricultural experiment station at Rothamsted, England, was the father of the randomized experiment, the general use of regression, and the mathematical derivation of the probability distributions of several important test statistics. JOAN FISHER BOX, *R. A. FISHER: THE LIFE OF A SCIENTIST* 113, 140 (1978). He was not the originator of tests of significance, but his writings on statistics in scientific research were exceedingly influential. *See, e.g.,* MOORE & MCCABE, *supra* note 149, at 425 (“One reason for the common use of . . . 0.05 is the great influence of Sir R. A. Fisher.”). Fisher wrote:

[I]t is convenient to draw the line at about the level at which we can say: “Either there is something in the treatment, or a coincidence has occurred such as does not occur more than once in twenty trials.” . . . If one in twenty does not seem high enough odds, we may, if we prefer it, draw the line at one in fifty (the 2 per cent. point), or one in a hundred (the 1 per cent. point). Personally, the writer prefers to set a low standard of significance at the 5 per cent. point, and ignore entirely all results which fail to reach that level. A scientific fact should be regarded as experimentally established only if a properly designed experiment rarely fails to give this level of significance.

R.A. Fisher, *The Arrangement of Field Experiments*. 33 J. MINISTRY AGRIC. OF GR. BRIT. 503, 504 (1926) (emphasis omitted). As one contemporary statistician has remarked: “There you have it. Fisher thought 5% was about right, and who was there to disagree with the master?” DAVID MOORE, *STATISTICS: CONCEPTS AND CONTROVERSIES* 292 (1979). For a more complete analysis of Fisher’s remarks, see Gerard E. Dallal, *Why P=0.05?*, <http://www.jerrydallal.com/LHSP/p05.htm> (last visited Dec. 22, 2008).

demanding) number. The essential point is that the significance level indicates the false-positive probability. Or rather, it is supposed to. There are ways to cheat.

Imagine an unscrupulous pharmaceutical company—the Evil Company—bent on “proving” that its newest miracle drug cures cancer. It commissions twenty randomized, controlled experiments with the drug. In each of these parallel clinical trials, the investigators, patients, or referring physicians think that they are participating in the only trial of the drug. In each trial, fifty patients are assigned to the control group and fifty to the treatment group. In the absence of the drug, the chance of remission is known to be 1/4. The twenty independent and isolated experiments are conducted. In one trial, there were nine more remissions in the treated patients than in the controls. All the other group differences were less than eight, and in all the trials combined, only twenty of the treated patients did better than the controls. Knowing all this, the hypothesis tester would conclude that even though the one reported trial is significant at the 0.05 level,¹⁵⁹ the *p*-value for all twenty trials combined is 0.30, which does not even approach significance.¹⁶⁰ The Evil Company, however, suppresses the nineteen negative findings and submits the one statistically significant finding to the Food and Drug Administration (“FDA”). A hoodwinked agency approves the drug.

In this hypothetical case, the nominal significance level of 0.05 greatly overstates the value of the positive finding when all twenty trials are considered. When we correct for the trawl through all the experiments, we realize that the evidence as a whole is not impressive. It would be somewhat surprising to find a significant difference in the first and only clinical trial if the drug were totally ineffective; however, it is not at all surprising to discover an apparent effect—a “hit”—in one or more of the twenty clinical trials. In fact,

159. Let X_1 be the observed number of remissions in the treatment group and X_2 be the number in the control group. Under the null hypothesis, X_1 and X_2 are binomially distributed with $n = 50$ and $\theta = 0.25$. Their difference $X_1 - X_2$ is approximately normal with mean 0 and standard deviation $\sigma = [2n\theta(1-\theta)]^{1/2} = 4.33$. If the alternative to the null hypothesis is that the drug could be either effective or detrimental, then any difference $|X_1 - X_2| > 1.96\sigma = 8.49$ is significant at the 0.05 level. If it is known that the drug cannot make things worse, then any difference of more than $1.64\sigma = 7.10$ in favor of the treatment group is significant at the 0.05 level.

160. Pooling the trials means that the standard deviation of the difference in the remissions is $[2(1000)(1/4)(3/4)]^{1/2} = 19.37$. Hence, $z = 20/19.37 = 1.03$. About 30% of the area under the standard normal curve lies in the corresponding tails. This is the two-tailed *p*-value, which is used if we cannot be sure in advance of the experiments that the drug can only be beneficial. The one-tailed *p*-value is 0.15, which also is not significant at the 0.05 level.

the Evil Company's strategy of presenting only the significant results is even more likely to produce a false positive than the pooled p -value of 0.30 suggests. The probability of at least one significant result in twenty can be computed in three easy steps: (1) the chance that there will *not* be a hit for a useless drug in each independent trial is $1 - .05$, or 95%; (2) the chance of no hit for all twenty trials is therefore $.95^{20} = .36$; and (3) the chance of one or more hits is the remaining probability, $1 - .36 = .64$. In other words, if we were to declare as effective those drugs that show an apparently significant effect at the 0.05 level—when, unbeknownst to us, the ineffective drugs are tested twenty times to find at least one such outcome—we would err 64% of the time. Because we have repeatedly “tested” the null hypothesis, the significance level used in the individual test is a misleading indicator of the false-positive error rate of our decision procedure. The true significance level is 0.64—a far cry from the nominal 0.05.

Of course, the Evil Company is but a dagger of the mind, a false creation.¹⁶¹ But this sort of thing happens inadvertently. It is known as publication bias.¹⁶² Not knowing of unpublished negative findings, researchers may be condemned to repeat other people's experiments until, at last, a “significant” difference emerges in a single experiment that sees the light of day.

Cognizant of this type of threat to conventional hypothesis tests, the NAS committees treated the random-match probability as though it were a p -value for the one clinical trial cherry-picked by the Evil Company. The 1992 committee worried that “[i]f a pattern has a frequency of 1 in 10,000, there would still be a considerable probability (about 10%) of seeing it by chance in a databank of 1,000 people.”¹⁶³ The 1996 committee offered a similar example:

[I]f we toss 20 reputedly unbiased coins once each, there is roughly one chance in a million that all 20 will show heads. According to standard statistical logic, the occurrence of this highly unlikely event would be regarded as evidence discrediting the hypothesis that the coins are unbiased. But if we repeat this experiment of 20 tosses a large enough number of times, there will be a high probability that all 20 coins will show heads in at least one experiment. In that case, an event of

161. Or is it? See *News in Brief: British Government to Demand Clinical Trial Data*, 452 NATURE 141, 141 (2008); Drummond Rennie, *When Evidence Isn't: Trials, Drug Companies and the FDA*, 15 J. L. & POL'Y 991, 1002–03 (2007).

162. Susan J. Landers, *Study Focuses on Publication Bias in Journals*, AM. MED. NEWS, Feb. 18, 2008, <http://www.ama-assn.org/amednews/2008/02/18/hlsb0218.htm>.

163. NRC I, *supra* note 15, at 124.

20 heads would not be unusual and would not in itself be judged as evidence that the coins are biased. The initial identification of a suspect through a search of a DNA database is analogous to performing the coin-toss experiment many times: A match by chance alone is more likely the larger the number of profiles examined.¹⁶⁴

The committees differed not over the need to adjust for the trawl, but only over which method for adjusting was more practical. As we saw in Part I, the 1992 report proposed retesting with new loci to get a new number (p_{rem}) for the random-match probability (the limit-the-loci approach), and the 1996 report described a statistical adjustment (multiplying by n) to the probability already on hand (the np rule).¹⁶⁵ Both methods are straightforward, and both increase the random-match probability p by many orders of magnitude, making the match seem less impressive to a jury.¹⁶⁶

Knowing the expected frequency of matches in an innocent database—one that does not include the true source of the crime-scene DNA—is useful for some purposes, and the value of np conveys a sense of how frequently trawls of innocent databases will produce matches. This untoward outcome has occurred at least once. In 2000, when the British database contained 660,000 DNA records, a Manchester man was linked to a burglary by a six-locus STR match.¹⁶⁷ The random-match probability was said to be one in thirty-seven

164. *Id.* at 134.

165. Although the 1996 report's analysis is more elaborate, the basic statistical reasoning behind its proposed correction is similar to our example of the twenty clinical trials of the useless anti-cancer drug. The idea is to consider how many times the apparently unlikely event had the chance to occur. Assume that the database of size n is composed entirely of unrelated people who have no involvement with the crime in question. Then the chance that the first individual trawled will match is just the random-match probability p . This also is the chance of a match on the second try, the third, and so on, for all n tries. The chance that any given individual will not match is therefore $1 - p$; the resulting chance that no one will match is $(1 - p)^n$. Hence, the chance that someone will match—even though all n are innocent—is $1 - (1 - p)^n$. This is the p -value adjusted for the trawl. The adjustment is the same as the one used in the example of the twenty clinical trials. If p is close to zero and n is very large, then $(1 - p)^n$ is approximately $1 - np$. Substituting this value in the previous expression $1 - (1 - p)^n$ establishes that the chance of at least one person in the database of innocent people matching the crime-scene sample is $1 - (1 - np) = np$. We have arrived at the committee's np rule.

166. The committee recommended multiplying instead of reducing the depth of the database trawl (by using fewer loci) to avoid the possibility that "[i]f the amount of DNA in the evidence sample is too small, following the recommendation in the 1992 report could leave too few additional loci for computing a match probability." NRC I, *supra* note 15, at 134.

167. Richard Willing, *Mismatch Calls DNA Tests into Question*, USA TODAY, Feb. 8, 2000, at 3A.

million, but the suspect had an excellent alibi.¹⁶⁸ When the STR typing was extended to ten loci, he was excluded.¹⁶⁹ Although a spokesman for the British database tried to explain that “[t]here’s no cause for alarm,” the case sent shock waves through the DNA testing community, which “fear[ed] that the error might prompt thousands of people convicted through DNA testing to appeal their convictions.”¹⁷⁰

In fact, the incident showed no error in either the DNA typing or the estimate of the random-match probability. What it did show was that p understates the chance that, sooner or later, an innocent person’s DNA will match. The product np is 66/3700, which is about 1/56. If the one-in-thirty-seven-million figure is typical of the six-locus random-match probability for the profiles in the database, then statistically one should expect a match about one time in fifty-six when the database (i.e., everyone in it) is innocent of the crime being investigated. The real surprise is that no six-locus matches to innocent people were seen in any of the hundreds or thousands of earlier trawls of the database. Yet, the manager of the FBI’s national DNA index system, NDIS, found the incident “mind-blowing.”¹⁷¹

The British experience shows that when considering how often database trawls will generate cold hits in databases that do not contain the true source, the frequentist perspective of the NRC committees makes sense. Even with perfect laboratory testing, unless the DNA type is unique, trawling innocent databases eventually will produce false matches more often than the random-match probability would suggest. But it does not follow that the existence of a trawl degrades the probative value of the evidence against the defendant. The next Section presents an intuitive explanation of this apparent paradox.

B. The Bayesian Rejoinder

To probe—and undermine—the intuition that a database trawl produces less powerful evidence than a confirmation match, consider two cases. In Case I, the defendant was identified through a trawl and further investigation produced confirmatory evidence. In Case II, the same confirmatory evidence was known at the outset, making the defendant a suspect. The police did not bother to secure a DNA sample from him, however, because they knew that his thirteen-locus

168. *Id.*

169. *Id.*

170. *Id.*

171. *Id.*

STR genotype was already included in the state's convicted-offender database. To be on the safe side, rather than just compare the crime-scene genotypes to the defendant's record in the database, they ordered a full search through the database. This search showed that the defendant matched and that no one else did. The only difference between the package of evidence in the two cases is the order in which it was uncovered.¹⁷² In Case I, the police trawled, then "confirmed." In Case II, they "confirmed," then trawled. It is hard to see why the evidence in Case I is any less probative than that in Case II.¹⁷³

In understanding the impact of a database trawl in a criminal investigation, these hypothetical cases are more salient than the chicanery of the Evil Company because they focus on the critical legal question about the evidence in a trawl case. The frequentist objection to the random-match probability in a trawl case is that the probability is a misleading, nominal p -value that should be replaced by an

172. In the trawl-first case there is the possibility that "the subsequently discovered evidence was tainted by suggestiveness, given that the DNA match motivated the police, and possibly witnesses as well, to confirm the suggestion that the perpetrator had been found." Donnelly & Friedman, *supra* note 12, at 959. The hypothetical in the text excludes this possibility. When it does occur, the nongenetic evidence is not identical in both cases. An eyewitness identification in a lineup tainted by knowledge of which person in the line-up has matching DNA is not the same as an eyewitness identification in a properly conducted lineup. This indicates that police should guard against tainting the nongenetic evidence in this fashion and that defense counsel should be alert to the possibility so that they can expose it at trial. Richard Lempert, *Some Caveats Concerning DNA as Criminal Identification Evidence: With Thanks to the Reverend Bayes*, 13 CARDOZO L. REV. 303, 319 (1991).

Another possible complication is that the subsequent investigation in the trawl-first case might produce more evidence than an ordinary investigation. Without the lead from the trawl, the police might have expended their limited resources in a more diffuse and less effective investigation and, hence, have acquired less adequate or complete nongenetic evidence. But this modification of the hypothetical case in the text does not mean that a match in the database is weaker evidence than a DNA match to an existing suspect. It simply means that the other evidence in the case might differ depending on when a DNA match is obtained, and it suggests that trawling first can be a good investigative strategy.

173. Professor Ronald Allen pointed out to me one way in which Case I could be more persuasive. It turns on the possibility that the databank does not contain DNA from the individual whose DNA was left at the crime-scene but does contain a sample from an identical twin. In these rare situations, the twin will emerge as a suspect from the database trawl, and if and when some confirmatory evidence is acquired, the investigation may end. On the other hand, if the police start by looking for nongenetic evidence, they are more likely to focus on the correct twin in the first place. In both cases, there could be a similar package of genetic and nongenetic evidence. This innocent-twin effect, however, is quite different from the frequentist concern that the evidence is degraded by the many opportunities for a match to an unrelated individual, and it is not an argument to replace the random-match probability with np or p_{rem} .

adjusted figure to account for the increased rate of false positives due to the search strategy. In more common parlance, the claim is that a jury that is not informed about the trawl and only hears about the incredibly small estimated frequency or random-match probability will overvalue the DNA match relative to a fully informed jury that knows about the trawl and the frequency of errors when searching large databases.¹⁷⁴ In the case of the Evil Company, the nominal p -value was misleading and an adjustment was required because the evidence was not simply the data in one controlled experiment. It was the data from that clinical trial plus nineteen others. One solution there would be to inform the FDA of all the results so that it can evaluate all the evidence about the drug, not just a selected part of it. As we saw with a few simple calculations, this would put the totality of evidence within the normal range expected under the null hypothesis.

When we ask about the impact of the additional evidence on a match from a DNA database trawl, however, we find that the effect is very different. This Section will show that a jury that learns that the match comes from the trawl should not regard the match as less significant than the same match in a confirmation case. Consequently, presenting the random-match probability and ignoring the fact that it came from a search is not prejudicial to the defendant. To establish this result, it is helpful to approach this question in two steps. First, we consider what the import of the DNA evidence would be if it consisted only of the one match between the defendant's DNA and the crime-scene sample (because he was the only person tested). Then, we compare the impact of the match when the data from the trawl are added to give the full picture. In the case of the Evil Company, the omitted evidence made it less probable that the drug was effective. In the database trawl case, however, we will find the opposite effect. If anything, the omitted evidence makes it more probable that the defendant is the source. On reflection, this result is entirely natural. When there is a trawl, the DNA evidence is more complete. It includes not only the fact that the defendant matches, but also the fact that other people were tested and did not match. The more people who are excluded, the more probable it is that any one of the remaining individuals—including the defendant—is the source. Compared to testing only the defendant, trawling a database

174. The reason not to disclose the trawl to the jury is that doing so could prejudice the defendant by revealing that he engaged in conduct that triggered his inclusion in a law enforcement database. Of course, when evidence of the defendant's prior crimes are admissible for other purposes, the concern evaporates.

that includes individuals who could have committed the crime therefore increases the probability that the defendant is the source. A database search is generally more probative than a single-suspect search.

The remainder of this Section applies some basic ideas from probability theory and presents numerical examples to buttress this intuition. The theory is Bayesian in that it conceives of an idealized juror's degree of belief in the proposition that the defendant left the crime-scene DNA as conforming to the laws of probability,¹⁷⁵ and it assumes that probability theory pertains to unique events and hypotheses about those events.¹⁷⁶ Despite the need to introduce a number of symbols for the sake of clarity and conciseness, no advanced mathematics is necessary. Basic algebra provides an informal proof of the claim that, contrary to the perception in the media of "a national problem" with "numbers that exaggerate the significance of DNA matches in 'cold hit' cases"¹⁷⁷ and the defense arguments in cases like *Puckett*,¹⁷⁸ a cold hit in a suitable database should be at least as convincing as a confirmation match. After defining some notation, I describe a more general version of Bayes' rule (a formula involving conditional probabilities) than is typically seen in the legal literature. I apply this formula to assess the impact of a database trawl on the central proposition in DNA cases—whether the defendant is the source of the crime-scene DNA.

Despite the frequentist insight that (for profiles that are not unique) a vast number of database trawls will lead to more DNA

175. See, e.g., Lempert, *supra* note 33, at 1035 (referring to an "ideal juror" in this way).

176. See, e.g., D.H. Kaye, *What is Bayesianism?*, in *PROBABILITY AND INFERENCE IN THE LAW OF EVIDENCE: THE USES AND LIMITS OF BAYESIANISM 1* (Peter Tillers & Eric D. Green eds., 1988). This Bayesian treatment is developed elsewhere more exhaustively for trace evidence (items or marks left at crime-scenes) generally, but the more rigorous presentations can be quite technical. E.g., A. P. Dawid & J. Mortera, *Coherent Analysis of Forensic Identification Evidence*, 58 J. ROYAL STAT. SOC'Y (SERIES B) 425 (1996) (analyzing with mathematical sophistication the effect of trace evidence on the proposition that an individual is the source of the evidence). For Bayesian analyses specific to DNA traces, see, for example, Balding, *supra* note 11, at 470–72, and see generally David J. Balding & Peter Donnelly, *Evaluating DNA Profile Evidence When the Suspect is Identified Through a Database Search*, 41 J. FORENSIC SCI. 603 (1996), and David J. Balding & Richard A. Nichols, *DNA Profile Match Probability Calculation: How to Allow for Population Stratification, Relatedness, Database Selection and Single Bands*, 64 FORENSIC SCI. INT'L 125 (1994). Expositions written for nonscientists include Balding, *supra* note 11, at 472, Donnelly & Friedman, *supra* note 12, and Walsh & Buckleton, *supra* note 22.

177. Felch & Dolan, *supra* note 4.

178. *Id.*

matches to innocent people than the random-match probability would suggest,¹⁷⁹ the Bayesian analysis shows that a match coming out of a trawl increases the probability that the defendant so identified is the source of the DNA in question at least as much as does a match to an individual who came to the attention of the police for other reasons. This finding demonstrates that the frequentist concern with multiple testing notwithstanding, the random-match probability is as good an indication of probative value in the trawl case as it is in a confirmation case. To facilitate the analysis, we start with some definitions.

1. The Ingredients

Bayesian analysis of evidence always uses five concepts: events, hypotheses, prior probabilities, likelihoods, and posterior probabilities. With DNA evidence, one critical event is that the defendant has the same DNA genotype as that associated with the crime. This matching individual can be labeled " I_0 ."¹⁸⁰ Of course, the defendant is not the only possibility; before considering the detailed evidence in a case, jurors must regard many other individuals in the general population as possible sources. The nature of this "suspect population" is discussed more fully in later Sections. At this point, we simply note that everyone in the relevant population of N people can be listed by a name I_j , where j goes from 0 to $N - 1$.¹⁸¹

Corresponding to each person is the hypothesis S_j , which states that I_j is the true source. If the relevant population is the entire world, then based on our background information about the world and any other (nongenetic) evidence in the case, every person has some probability $\Pr(S_j) = \pi_j$ of being the source. This is a "prior probability." It applies before we consider the DNA evidence. Individuals who are on the other side of the globe or the wrong sex

179. *Id.* (quoting Stanford mathematician, Keith Devlin, for the view that "[i]t is only a matter of time until someone is wrongfully convicted because of this").

180. I use this odd name because such nomenclature leads to a compact way to refer to everyone in the population of N possible suspects (which could consist of almost everyone in the world). More than one individual might match, especially if the crime scene contains so little DNA that only a few alleles can be detected. Such situations are not analyzed here.

181. The capital N , which stands for population size, should not be confused with n , which is the database size. Indeed, we can think of the confirmation case as a degenerate database case in which a "database" of size $n = 1$ consists of the suspect's DNA profile. This also reminds us that the issue is not limited to trawls through pre-established databases. In principle, it arises every time investigators test and look at the DNA types of more than one individual to compare them to the crime-scene DNA profile.

would have prior probabilities near zero.¹⁸² An individual selected for DNA matching based on other, independent evidence should have a higher prior probability than someone who is tested for no other reason than that his DNA profile is on file.

So far, we have enumerated the hypotheses of interest—the ones that answer the question, “Who is the source of the crime-scene DNA?” Now we refer to the events that have occurred as E .¹⁸³ The DNA evidence E in a pure confirmation case (only one person’s DNA has been compared to the crime-scene sample) is the match to I_0 . This event (denoted as $E_1 = M_0$)¹⁸⁴ has a probability of arising under each of the N hypotheses. More generally, the evidence E is E_n , where n is the size of the database that has been searched. For $n > 1$, this evidence consists of the match to defendant and the exclusion of everyone else in the database: $E_n = M_0 \& X_1 \& \dots \& X_{n-1}$.

Colloquially, “likelihood” is synonymous with “probability” but it has a more specialized meaning here. For a given item or body of evidence, the likelihood varies across the hypotheses S_j . When evidence is more probable under one hypothesis than another, the hypothesis that makes the evidence more probable is said to have a greater “likelihood.” The likelihood L_0 involving the match to the defendant is the probability of the DNA data computed on the assumption that the defendant is the source. Since the defendant is sure to have the same genotypes if he is the source, the probability of the evidence E_1 given S_0 in the simple confirmation case is $L_0 = \Pr(E_1|S_0) = 1$.¹⁸⁵ The likelihood for each hypothesis is $L_j = \Pr(E_1|S_j)$.¹⁸⁶

182. For simplicity, it will be convenient to treat individuals who apparently could not have committed the crime (astronauts in orbit at the relevant time, musicians who were performing onstage, and so on) as if their prior probability is zero. This reduces the size N of the initial population. The purist alternative is to keep the population extremely inclusive but give many people miniscule prior probabilities that reflect such unlikely possibilities as a plot to place a look-alike of the astronaut in orbit.

183. I use E interchangeably with “evidence,” although technically, the evidence in court is a report about these events. Such reports can be in error, and a more detailed study would model this possibility. Inasmuch as this complication does not affect the conclusions here, however, I shall put it to the side.

184. M_0 simply refers to a match, (M) to I_0 . An exclusion (X) of I_0 would be denoted X_0 .

185. This assumes that the laboratory has correctly identified all the genotypes in both samples.

186. More precisely, the likelihood is defined so as to be proportional to the conditional probability. Because the proportionality constant cancels out of the equations needed here, we can ignore it.

There is a logical difference between the likelihood $L_0 = \Pr(E_1|S_0)$ and its transpose, $\Pr(S_0|E_1)$.¹⁸⁷ The transposed probability is called a “posterior probability,” because it arises after the evidence E_1 is obtained. Evidence tending to establish a defendant’s guilt increases the prior probability of guilt to a higher level, while evidence tending to show innocence produces a lower posterior probability. The posterior probability that any individual I_j is the source is abbreviated as $\pi_j' = \Pr(S_j|E_1)$. Thus, the posterior probability that I_0 is the source can be written as $\pi_0' = \Pr(S_0|E_1)$. More generally (for a trawl through a database of size n), this posterior probability is $\pi_0' = \Pr(S_0|E_n)$.

These five ingredients—events, hypotheses, prior probabilities, likelihoods, and posterior probabilities—are interrelated. Bayes’ rule is a recipe for combining them to arrive at the posterior probability that any individual is the source of the crime-scene DNA. In most legal writing, a version that uses only two hypotheses is presented.¹⁸⁸ We need to be more comprehensive to solve the database-trawl puzzle, for there are many hypotheses S_j to assess. Even so, Bayes’ rule remains fairly simple.

2. The Recipes

Bayes’ rule expresses how an item of evidence shifts the prior probabilities assigned to a set of hypotheses. The formula instructs us to weight each prior probability by its likelihood and to combine these weighted prior probabilities to obtain the posterior probabilities. An example will convey the basic idea. Suppose we take all the hearts from a standard deck of playing cards and shuffle them well. We draw one card and give it to a friend to view. What is the probability that it is a king? I will abbreviate this possibility as C_K , for “card is a king.” Since there is one king in the reduced deck of thirteen hearts and it is as likely to be drawn as any other card, the prior probability is $\Pr(C_K) = 1/13$. Now we learn something more about the card—our friend (who never errs or lies) assures us that it is a picture card. This evidence E raises the probability. The revised probability is $\Pr(C_K|E)$, the conditional probability that the card is a king given that it is a picture card. It is just the proportion of picture

187. In the special case of $\pi_0 = 1/2$ and small $\Pr(E_1|S_0)$, the naïve transposition is approximately correct. A.P. Dawid, *The Island Problem: Coherent Use of Identification Evidence*, in ASPECTS OF UNCERTAINTY: A TRIBUTE TO D. V. LINDLEY 159, 160 (P.R. Freedman & A.F.M. Smith eds., 1994).

188. See, e.g., Finkelstein & Fairley, *supra* note 34, at 498; David H. Kaye, *DNA Evidence: Probability, Population Genetics and the Courts*, 7 HARV. J.L. & TECH. 101, 165 (1993); Tribe, *supra* note 34, at 1353.

cards that are kings, which is $1/3$. The evidence has raised the prior probability $\pi_K = 1/13$ to the posterior probability $\pi_K' = 1/3$.

This card problem is simple enough that we can compute the posterior probability directly (by counting the number of picture cards). In more complicated cases, we can use Bayes' rule written out in terms of the prior probabilities and likelihoods for all the cards. The rule states that the posterior probability of the hypothesis of interest is

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood for hypothesis of interest}}{\text{sum of all (priors} \times \text{likelihoods)}} \quad (1)$$

In the card problem, the probability that a king has been drawn given that a picture card has been drawn is (a) the likelihood for a king weighted by the prior probability divided by (b) a weighted sum of the likelihoods for all the cards (the weights being the prior probabilities).

To verify that the rule works, we need to know the priors and the likelihoods. At the outset, every card has the same chance of being picked. Since all the priors have the same value, they cancel out of equation (1). The equation thus simplifies to $\pi_i' = L_K / [L_1 + \dots + L_{10} + L_{\text{Jack}} + L_{\text{Queen}} + L_K]$. The likelihood of the hypothesis that the card is a king is the probability of the evidence (a picture card) given the hypothesis that it is a king: $L_K = \Pr(E|C_K) = 1$. The likelihood for the two other picture cards, a jack and a queen, is the same. Because cards one through ten cannot produce the report that they are a picture card, their likelihoods are zero. We conclude that $\pi_K' = 1 / (0 + \dots + 0 + 1 + 1 + 1) = 1/3$.

That Bayes' rule works in the card example is no surprise since the formula is easily derived from the axioms and definitions of probability theory. Furthermore, Bayes' theorem is uncontroversial in this trivial card game because objective probabilities can be attached to the events or hypotheses. But the rule also can be applied to subjective probabilities¹⁸⁹ or to degrees of belief about propositions of fact.¹⁹⁰

189. LEONARD J. SAVAGE, *THE FOUNDATIONS OF STATISTICS* 30 (2d ed., Dover Publications 1972) (1954); Peter C. Fishburn, *The Axioms of Subjective Probability*, 1 *STAT. SCI.* 335, 335 (1986).

190. RICHARD C. JEFFREY, *THE LOGIC OF DECISION* 184-86 (2d ed. 1983); D.V. LINDLEY, *MAKING DECISIONS* 45-46 (2d ed. 1985). Bayesian statisticians are more eclectic than their frequentist colleagues in the sense that they are willing to extend the domain of probability beyond events that are governed by physical random processes. They do not claim that the personal beliefs of real people actually conform to the

A version of Bayes' rule for a DNA match is

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood for defendant}}{\text{prior} \times \text{likelihood for defendant} + \text{sum of all (priors} \times \text{likelihoods for everybody else)}}. \quad (2)$$

The denominator contains all the weighted likelihoods, but they are broken into two groups—the one for the defendant and those for everybody else. The formula applies to trawl cases and confirmation cases alike.

3. The Dishes

We now show that learning about the results of the database search as opposed to learning only that the defendant matches does not lower the posterior probability that the defendant is the source. Only a verbal sketch of the mathematical reasoning is given here. A more explicit derivation is provided in the appendix.

We begin by putting the likelihoods in the denominator in a useful order. The weighted likelihood for the defendant goes first, followed by the terms for everyone else in the database, followed by the terms from the untested portion of the population. The posterior probability becomes:

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood for suspect}}{\left\{ \begin{array}{l} \text{prior} \times \text{likelihood for suspect} \\ + \text{sum of (priors} \times \text{likelihoods for everyone else in database)} \\ + \text{sum of (priors} \times \text{likelihoods for untested individuals)} \end{array} \right\}}. \quad (3)$$

The crucial observation about (3) is that the sum of the weighted priors for everyone in the database drops out of the equation.¹⁹¹ The

probability calculus anymore than accountants claim that everyone adds up numbers correctly. Rather, they maintain that people should follow Bayes' rule in articulating their degrees of beliefs—just as they should follow the rules of arithmetic in adding numbers or the rules of deductive logic in evaluating the truth of propositions. Thus, one argument for treating degrees of belief as conforming to the axioms of probability theory (and hence, Bayes' rule) is that only this procedure assigns the same degrees of belief to logically equivalent propositions. For instance, only degrees of belief that act like probabilities will produce mutually consistent assessments of propositions, such as: "It is not the case that both A and B are true," and, "Either not-A is true or not-B is true." See generally, BRIAN SKYRMS, *CHOICE AND CHANCE: AN INTRODUCTION TO INDUCTIVE LOGIC* (4th ed. 2000).

191. This setup assumes that the database is nested inside the suspect population. Real databases include samples from offenders who are behind bars and could not be the perpetrators of new crimes in the outside world. They may even include individuals who

trawl tells us that, within the database, only the defendant has the matching genotype. Some untested individuals might share that genotype, but everyone else in the database has been excluded. Therefore, the probability that a tested individual I_j would have genotypes that are inconsistent with the crime-scene stain when this person is the source is $L_j = 0$. As such, (3) simplifies to:

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood for suspect}}{\left\{ \begin{array}{l} \text{prior} \times \text{likelihood for suspect} \\ + \text{sum of (priors} \times \text{likelihoods for untested individuals)} \end{array} \right\}}. \quad (4)$$

We are ready to compare the posterior probability of S_0 knowing that there has been a database search to the probability based on the mistaken assumption that there has been no trawl (and everyone else in the population is untested). If the priors are the same in both situations, the only thing that changes in (4) is the number of weighted likelihoods that are added up in the sum for the untested individuals. Searching a database reduces the number of untested individuals. For instance, a trawl through a database of size one hundred that includes the defendant moves ninety-nine individuals from the “untested” category into the “database” category—where they are excluded. Their weighted likelihoods disappear from the denominator. A smaller denominator means a bigger fraction. Because of this “denominator effect,” the trawl increases the posterior probability even more than a confirmation match to a single suspect.

The denominator effect is miniscule when the database is a tiny part of the population. In the limit, as the database grows to encompass the whole population, however, the sum of the weighted likelihoods from outside the database goes to zero. All that remains in the denominator is the weighted likelihood for the defendant. The posterior probability becomes one—just as we would expect when everyone in the population except for the defendant has been excluded.

Based on this reasoning about the effect of excluding everyone except for the defendant in a large database, it seems perverse to

are dead but whose records have not been purged. When evidence from a decades-old crime is tested, young offenders who could not possibly have committed the crime will be in the database. Mathematically, this is not a problem. One can define the suspect population very liberally and assign prior probabilities that are approximately zero to the unrealistic “suspects.” Alternatively, one can drop these people from the “suspect population” and reduce the offender database size n accordingly.

inflate the random-match probability. When one considers the actual DNA evidence in the trawl case—that the defendant matches *and no one else does*—we can be *more* confident that the defendant, to the exclusion of everyone else in the world, is the source of the stain. Consequently, “[n]o downward adjustment in the force of the evidence is appropriate.”¹⁹²

C. Reconciling Conflicting Frequentist and Bayesian Intuitions

The discussion so far has shown that the extended search for a significant result degrades the probative value of the “significant” finding in the cancer-drug case but not, it seems, in the hit in the DNA database trawl. Several factors underlie this divergence. Clearly, it is wrong for the Evil Company to search for significance and then to present the one instance of a small *p*-value as if there had been no search. The reason is simple: the omitted information contradicts the company’s claim of therapeutic effectiveness. In contrast, if all the other clinical trials were consistent with the company’s claim, the failure to mention them would not prejudice the case for approving the drug. The trawl case is similar to a series of successful clinical trials, all rejecting the null hypothesis for the nominal *p*-values. The additional evidence—that everyone else in the database is excluded—is consistent with the claim of defendant’s guilt. The lack of other hits in the database trawl therefore has an effect opposite to that of the lack of other significant differences in the clinical trials. It supplies compatible rather than contradictory data. There is no paradox in stating that ascertainment bias is a problem when testing drugs but not when looking for matches for identification in a database.

Likewise, the database trawl differs from the coin-tossing example in the 1996 NRC report. There, the sequences of twenty tosses that do not yield streaks of heads undercut the hypothesis that the coin is biased toward heads. A more apt analogy to a DNA database trawl would be a search of a large room that contains fair

192. BALDING *supra* note 148, at 36; Donnelly & Friedman, *supra* note 12, at 933; authorities cited, *supra* note 176. Indeed, the denominator effect is not the only thing that increases the posterior probability. Our exposition treated the random-match probability as a known quantity not affected by the testing in the instant case. The preceding authorities also argue that the very failure to find additional matches in the database increases the probability that the profile in question is truly rare, giving a second reason to be more confident in a trawl case than in a confirmation one. Like the denominator effect, however, this factor has only a small impact on the posterior probability.

coins together with a few two-headed coins.¹⁹³ We draw a single coin, toss it twenty times, and observe twenty heads. This presents powerful evidence that the coin is two-headed. The result is certain if the coin is two-headed, and it is almost impossible when the coin is fair. The likelihood for the fair coin is $p = 1/2^{20}$, or about one in a million; the “likelihood ratio” is then $1/(1/2^{20}) = 2^{20}$, or about a million. The data of twenty heads are a million times more likely to arise when the coin is two-headed than when it is a fair coin.¹⁹⁴

Now suppose that instead of tossing a single coin, we sample 1001 coins, toss each of them twenty times, and find that just one exhibits twenty heads. The trawl through the coins does not make the hypothesis that the coin is two-sided less probable. To the contrary, excluding one thousand coins makes it *more* probable that the coin is one of the remaining double-headed ones. If we fail to disclose that we found the coin by testing many coins instead of one, we would not be suppressing any data that undermine the contention that the coin is two-headed. Instead, we would be understating the case for the two-headed coin.

Yet, many people have the sense that the one-hit database search is less diagnostic of identity than the one-hit confirmation test. In fact, one article goes so far as to claim that “a Bayesian analysis suggests that this evidence [from a trawl] has no probative value.”¹⁹⁵ According to the article’s authors, “a match between the suspect and perpetrator is nearly certain if the suspect is in fact guilty. But a match between the suspect and perpetrator is also certain if he is not guilty, because he was chosen based on the fact he matches.”¹⁹⁶ The last statement, however, is confused or at least confusing. It seems to be addressing the hypothesis that *some unspecified individual* in the database is guilty rather than the hypothesis that *a named defendant* in the database is guilty. Before the database search is conducted, it is not certain that anyone in the database will match, let alone that a specific, named individual—call him Mr. J—will match. Naturally, an innocent person is likely to match if the database is large enough, if the actual source is not in it, and if the profile is not unique to this

193. Cf. Donnelly & Friedman, *supra* note 12, at 950–51 (analyzing a similar problem involving coins).

194. This likelihood ratio is an excellent measure of the probative value of evidence with respect to two simple and complementary hypotheses like the ones in this example. See, e.g., 1 MCCORMICK, *supra* note 96, at 730–32; Kaye & Koehler, *supra* note 108, at 648–50.

195. William C. Thompson & Simon Ford, *DNA Typing: Acceptance and Weight of the New Genetic Identification Tests*, 75 VA. L. REV. 45, 101 (1989).

196. *Id.*

individual. These conditions are what give rise to the valid frequentist concern that, sooner or later, database trawls for nonunique profiles will produce false hits.¹⁹⁷

The Bayesian response is that the individual, not the database, is on trial.¹⁹⁸ The appropriate analysis focuses on the hypotheses S_j that “Mr. J is the source of the crime-scene sample.” Before the search, the probability that a given J is the source will be low (if there is no other reason to suspect him) or high (if there is other evidence pointing to him). In either case, these prior odds will increase dramatically as a result of the discovery that out of all the samples in the databank only Mr. J’s matches.¹⁹⁹ Therefore, the database finding has considerable probative value.

Yet, the defendant’s name is only attached after the search is over. Had Mr. K matched instead of Mr. J, he would have become the suspect in the case. So the feeling that the trawl has not proved much lingers. This doubt about probative value might emanate from the fact that the posterior probability in a “naked database match” case could be small. By a “naked match case,” I mean a case like the one noted in the introduction, where John Davis was prosecuted in San Francisco without any other evidence of guilt (besides such general facts as his residence and age).²⁰⁰

197. The FBI’s DNA Advisory Board “continue[s] to endorse the recommendation of the NRC II Report for the evaluation of evidence from a database search” because it finds “the LR [likelihood ratio] espoused by Balding and Donnelly (1996)” to be “unsatisfactory” as applied to an example involving a series of genotypes that have a frequency of 1/100,000. DNA Advisory Board, *supra* note 22. The example, however, merely establishes the same point—sooner or later, there will be false hits in a large database.

198. See *People v. Nelson*, 185 P.3d 49, 66 (Cal. 2008) (“[T]he database is not on trial. Only the defendant is.”) (internal quotation marks omitted) (citing 4 DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 32.11 at 110, 118–19 (2006)).

199. An appendix in the 1996 NRC report distinguishes between the two hypotheses and shows that the probability that a named individual is the source is larger following a database trawl. NRC II, *supra* note 16, at 163–65. The committee offered the *np* adjustment when considering the null hypothesis that no one in the database is the source versus the alternative that someone in the database is the source. See *id.* at 134–35 (“Suppose that we hypothesize that the evidence sample was not left by someone whose DNA profile is in the database (or a close relative of such a person) and find that under this hypothesis $P(M)$ is small. The usual statistical logic then leads to rejection of that hypothesis in favor of the alternative that (one of) the matching profile(s) in the database comes from the person who left the evidence sample.”); *id.* at 161 (“If one wishes to describe the impact of the DNA evidence under the hypothesis that the source of the evidence sample is someone in the database, then the likelihood ratio should be divided by $[n]$.”).

200. See *supra* notes 6–7 and accompanying text.

The idea that the database search is less informative because defendant was not already a suspect is misguided for two reasons. It confuses probative value—which relates to a shift in the odds—with the posterior odds—the odds after all the evidence is in.²⁰¹ In addition, it compares the posterior in a case of limited evidence to the posterior in a case of more extensive evidence. Naturally, the former will be less than the latter, but this does not mean that limited evidence is not probative and can be ignored. Rather, it means that, when the investigation is over, there may be too little evidence to warrant a conviction. Strong evidence is like a good play in football. It moves the ball a long way down the field. But if a team takes possession deep in its own territory, even a great play may not score a goal. The failure to score with one item of evidence hardly means that the evidence is weak or prejudicial. A DNA match in a trawl case—even a naked trawl case—is no less probative than a DNA match in a confirmation case.

To erase any residual doubt on this score, the next Section explicitly compares the probative force of both confirmation and trawl matches in several examples. The examples, and the more general reasoning behind them (which is found in the Appendix), show that trawling does not degrade the value of a match.

D. The Naked Trawl and Other Cases

A study of some examples of the impact of trawl evidence in various situations clarifies the nature of the naked-trawl case. This Section considers four different situations: (1) a pure confirmation case (nongenetic information points to the defendant, he is the only one tested, and he matches); (2) an investigate-then-trawl case (the nongenetic information is followed by a confirming database trawl that points only to the defendant); (3a) a trawl-then-investigate-successfully case (the trawl points only to the defendant and is followed by an investigation that uncovers the same confirming nongenetic information); and (3b) a trawl-then-investigate-unsuccessfully case (the trawl points only to the defendant and is followed by an investigation that uncovers no confirming nongenetic information).

I shall use a numerical example together with the physical metaphor of a mass, or fixed quantity, of probability that must be divided up among the hypotheses about the true source. We want to know how much of the mass ends up on the defendant in each

201. Kaye & Koehler, *supra* note 108, at 654–55.

situation. The example will illustrate how cases (2) and (3a) produce the strongest body of evidence. Case (1) ranks next, and case (3b), which has the least total evidence, comes in last. In every case, however, the trawl is equally probative, again establishing that there is no need to adjust the match probability that applies in the pure confirmation case.

1. The Pure Confirmation Case: $n = 1$

When eight-year-old Alicia Wade was abducted from her bedroom in San Diego, molested, then returned there, she did not tell anyone. The assault only became known after she complained to her mother that it hurt when she went to bathroom.²⁰² Child Protective Services placed Alicia in foster care.²⁰³ More than a year later, she told her psychotherapist that her father, James Wade, had attacked her.²⁰⁴ He was charged with lewd and lascivious acts causing great bodily injury.²⁰⁵ After reading the records of the therapy sessions obtained by James's defense counsel, however, the prosecutor "was convinced that they showed Alicia had been brainwashed" by the therapist.²⁰⁶ He turned to the crime lab to see if it could find something to confirm his impression. Using new equipment, a criminalist discovered a semen stain on the girl's nightshirt that had not been detected earlier. DNA testing excluded James as the source of the semen.²⁰⁷

The criminal case now centered on Albert Carder, who by then was in prison for assaulting other girls in the same neighborhood.²⁰⁸ DNA testing "showed that Carder could be the person who left the stain" and that the random-match probability was "less than one in three million."²⁰⁹ Before the DNA test was conducted police had

202. GEORGE "WOODY" CLARKE, JUSTICE AND SCIENCE: TRIALS AND TRIUMPHS OF DNA EVIDENCE 64 (2007).

203. *Id.* at 67–68.

204. *Id.* at 70.

205. *Id.*

206. *Id.* at 72.

207. Although the judge made a factual finding that James was innocent when the prosecutor's office moved for its charges to be dismissed, CPS persisted in its efforts to terminate parental custody. *Id.* at 74. Alicia's mother attempted suicide. John Wilkens & Jim Okerblom, *Rape-Case Therapist Gives Up License: Accused of Coercing Wade Girl's Allegation Against Dad*, SAN DIEGO UNION-TRIB. Mar. 20, 1996, at B1. The parents ultimately prevailed in juvenile court. *Id.* The therapist lost her license and paid one million dollars toward the settlement of the family's tort action against her and others involved in the case. *Id.*

208. CLARKE, *supra* note 202, at 69–70, 74.

209. *Id.* at 74.

reasons to believe that Carder was the culprit, but we may assume that this evidence was far from conclusive. Imagine that a juror hearing this nongenetic evidence starts with a prior probability of $1/1001$ that Carder is the source of the semen on the nightshirt. The total prior probability is $1 = 1001/1001$, so the remaining probability mass of $1000/1001$ is distributed in some fashion over everyone else in the population.

What about the likelihoods? The DNA evidence is that one person, Carder, has been tested, and he matches: $E_1 = M_0$. If Carder is the source then his genotype must match the genotype in the crime-scene sample. So the likelihood for Carder is just $\Pr(M_0|S_0) = 1$. Hence, the weighted likelihood for Carder is $1/1001 \times 1$. The likelihood for every other hypothesis is the chance that Carder would match if each untested man—and not Carder—were the source. For every unrelated man, this probability is the random-match probability p . Different amounts of the prior probability mass of $1000/1001$ may fall on each of these men, but the sum of the weighted likelihoods has to be $1000p/1001$.

Substituting this value into Bayes' rule (2), we find that the posterior is

$$\pi'_0 = \frac{\frac{1}{1001}}{\frac{1}{1001} + \frac{1000p}{1001}} = \frac{1}{1 + 1000p}. \quad (5)$$

Because the random-match probability p in the case was about one in three million, we conclude that $\pi'_0 = 1 / (1 + 1/3000) = 3000/3001 = 0.9997$. Thus, the fact that Carder's profile matches makes most of the prior probability assigned to everyone other than Carder flow to him, swamping his initially low prior probability. Unless one can explain the match on some other basis, Carder is almost certainly the source.

2. The Trawl-as-confirmation Case: $n > 1$

When the database is larger and only the defendant matches, we have more information about some of the people who were not tested in the pure confirmation case. They have been excluded. As people are excluded, their probability mass flows to the remaining individuals—including the defendant. The case against the defendant becomes even stronger. Suppose, for instance, that having identified Carder by nongenetic evidence, a San Diego County database of 100,000 people, including Carder, is searched. Again, the prior

probability for Carder is $\pi_0 = 1/1001$. Before the database trawl, the probability mass outside of Carder (1000/1001) is divided somehow among the men outside the database. Let's say that 1/6 of the prior probability of 1000/1001 falls on the convicted offenders in the database and 5/6 falls on the rest of the population (other than Carder). The latter group of untested men thus has the prior probability of (5/6) (1000/1001). Carder still has only a 1/1001 chance of being the source. The trawl eliminates the 99,999 other men in the database. Their probability mass flows to the defendant and to the many untested men. Carder's probability gets a small boost from this, and he gets a large chunk of the probability mass from the fact that he does match. Specifically, Bayes' rule reveals that

$$\pi'_0 = \frac{1/1001}{1/1001 + \frac{5}{6} [1000p/1001]} = \frac{6}{6 + 5000p}. \quad (6)$$

Since $p = 1/3,000,000$, Carder ends up with a total posterior of 18000/18005. This is marginally larger than the probability of 3000/3001 = 18000/18006 in the simple confirmation case. Here, using the unadjusted random-match probability is an excellent approximation to the correct value. There is no need to inflate the random-match probability in the investigate-then-trawl case.

It is not clear that the frequentists dispute this conclusion. The 1996 NRC recommendation was directed only to the "situation . . . in which the suspect is initially identified by searching a database to find a DNA profile matching that left at a crime scene."²¹⁰ When the defendant has been identified by nongenetic evidence, the frequentist concern with "searching for significance" does not apply.²¹¹ The defendant has not come to the government's attention by virtue of a cold hit in a database. He was a hot prospect; therefore, even from

210. NRC II, *supra* note 16, at 134.

211. The Bayesian statisticians who pioneered this kind of analysis sometimes use the trawl-as-confirmation case to argue that the frequentist demand for adjusting the random-match probability is misguided. Balding & Donnelly, *supra* note 176, at 605, write that the NRC recommendation, or at least "the rationale behind [it] . . . could lead to the rather absurd situation in which a cunning defense lawyer could insist on one and subsequently claim that the failure to find additional matches had substantially weakened the case against their client." See IAN EVETT & BRUCE WEIR, *INTERPRETING DNA EVIDENCE: STATISTICAL GENETICS FOR FORENSIC SCIENTISTS* 222 (1998). However, the 1996 recommendation to multiply p by the database size expressly applies only "[w]hen the suspect is found by a search of DNA databases." NRC II, *supra* note 16, at 161. The 1992 NRC report is less clear on this point. It refers to "the selection bias that is inherent in searching a databank" and broadly condemns "fish[ing] through the databank, trying out many hypotheses." NRC I, *supra* note 15, at 124.

the frequentist perspective, the fact that other profiles were compared to the one from the crime scene does not reduce the value of the match to the defendant. It enhances it.

3. The Trawl-then-confirm-successfully Case

All that remains are the cases in which the trawl comes first rather than last. These resemble data mining in that the government searches everywhere in the database to find one gold nugget to exhibit. Assume then that the gold nugget is Carder, who was unknown to investigators before the database trawl. With Carder's name at their disposal due to the trawl, investigators uncover the same clues that would have led them to Carder in the simple confirmation case, and so he goes on trial.

As suggested earlier, the order in which jurors process new information makes no difference to the final outcome.²¹² At the outset, a juror assigns some initial probability to each competing hypothesis.²¹³ Suppose that there are two items of evidence, E_A and E_B , to consider. At trial, each side generally can present its evidence in the order it chooses. If E_A precedes E_B , then a first application of Bayes' rule with likelihoods such as $\Pr(E_A|S_0)$ modifies the prior probability. This posterior probability $\Pr(S_0|E_A)$ becomes the prior probability in a second application of Bayes' rule with E_B as the evidence. Because multiplication is commutative (the product of any two numbers x and y is the same as the product of y and x), the resulting posterior probability is identical to what is obtained by starting with E_B , then considering E_A .²¹⁴ It also is the same as suspending judgment until both E_A and E_B are presented, then applying Bayes' rule only once with likelihoods such as $\Pr[(E_A \& E_B)|S_0]$. Thus, one can start with the nongenetic evidence or with the DNA evidence; the final probability, according to Bayes' theorem, will be the same.

Not only does the formula show that order is irrelevant, but intuitively it also seems that order should be irrelevant. An eyewitness might pick a defendant from a set of mug shots at the outset of an investigation, leading to further evidence against the individual. This is a trawl-first case (for a database of photographs). Alternatively, the witness might pick the defendant from mug shots after the police had already investigated the suspect. This is a trawl-

212. See *supra* notes 172–73 and accompanying text.

213. We will come back to what these might be later. See *infra* Part III.B.

214. See *infra* Appendix B.3.

last case. If the witness has not been influenced by the police suspicions in the trawl-last case, the match between the witness's memory and the photograph of the defendant is equally probative in both cases.²¹⁵ Consequently, no new calculation is needed to show that the trawl-then-investigate case can be handled the same way as the investigate-then-trawl case.²¹⁶

4. The Trawl-then-confirm-unsuccessfully Case

Finally, suppose that the police were unable to locate substantial nongenetic evidence against Carder. This is the naked-trawl case. Now, we can say only that Carder lived in San Diego when the abduction occurred and that he is a man in the right age range with matching DNA. The prior probability that Carder is the source cannot be 1/1001 as it was in the other cases. That figure was predicated on some nongenetic evidence of guilt that is not present in the naked trawl situation. Suppose we start with the much smaller prior of 1/1,000,001, trawl through a database of 100,000, and discover that only the defendant matches. As before, we stipulate that 5/6 of the prior probability (which is now 1,000,000/1,000,001) falls on the men outside the database. The posterior probability will be

$$\pi'_0 = \frac{1/1,000,001}{1/1,000,001 + \frac{5}{6} [1,000,000p/1,000,001]} = \frac{6}{6 + 5,000,000p}. \quad (7)$$

For $p = 1/3,000,000$, the posterior is only $18/23 = 0.78$, indicating that the total evidence is much less compelling in the naked-trawl case.

But this is not because *the trawl* is any less probative. The case as a whole is weaker because there is no incriminating nongenetic evidence to consider in formulating the prior probability. Suppose that instead of locating Carder by a DNA trawl, a detective dreamt that a person named Albert Carder was responsible for an assault, acquired a sample of Carder's DNA, learned that it matched the crime-scene sample, and could find no other proof of Carder's involvement. To apply Bayes' theorem, we simply adjust the prior probability of 1/1,000,001 according to the version of Bayes' rule that applies to a simple match case to obtain a posterior of

215. Cf. Donald A. Berry, *Comment*, 9 STAT. SCI. 252, 254 (1994) (stating that in both cases the "hypothesis generation and confirmation" should be treated symmetrically).

216. The Appendix gives the formulas that would permit a new calculation should one be desired.

$$\pi'_0 = \frac{1/1,000,001}{1/1,000,001 + 1,000,000p/1,000,001} = \frac{1}{1 + 1,000,000p}. \quad (8)$$

Since $p = 1/3,000,000$, the numerical value is $1 / (1 + 1/3) = 3/4 = 0.75$. This quantity is less the 0.78 for the naked-trawl case. Again, the trawl produces more powerful evidence than a simple match.

The crucial point is that E_n always has a bigger impact on the prior probability than does E_1 . In itself, a trawl that points only to the defendant is always more convincing evidence of guilt than the same match in the absence of the more extensive testing provided by the trawl. Naturally, if the prior probability, which is based on the non-DNA evidence, is smaller (as it is in a naked trawl case), then the posterior probability will be smaller. But the existence of a trawl is not what makes the totality of the evidence less convincing in a naked trawl case.

These findings refute the frequentist argument about multiple hypothesis testing. That argument implies that by diminishing the significance of the match, identifying the defendant by trawling weakens the total evidence against him. Yet, the DNA evidence is generally stronger due to the trawl because there is more of it, and all of it is consistent with the hypothesis S_0 that the defendant's DNA was left at the crime scene.

E. Back to Basics

Although the frequentist concern with data mining seems misplaced, it is worth reexamining the crucial premises of the Bayesian analysis. First, it assumes that the extent to which evidence alters the prior odds is what makes it probative.²¹⁷ Second, it holds that the order of the evidence is irrelevant.²¹⁸ Finally, the argument assumes that a prior probability that does not involve the evidence in question exists and is the same in a trawl-then-investigate case as it is

217. *E.g.*, Kaye & Koehler, *supra* note 108, at 646-47.

218. *E.g.*, John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1085 (1968) ("Since Bayes' Theorem is applied again and again by successive multiplications, . . . the order in which evidence comes in will not affect our rational decisionmaker though differences of emphasis caused by differing orders of proof may, of course, have a psychological effect on a real-life jury."); Peter Tillers, *Webs of Things in the Mind: A New Science of Evidence*, 87 MICH. L. REV. 1225, 1249 (1989) ("a conventional Bayesian [does not believe] that the order in which evidence is presented is probatively significant"); *supra* notes 212-14 and accompanying text.

in an investigate-then-trawl case.²¹⁹ These three premises are so clear to Bayesians that the preceding argument (usually presented more elliptically or compactly) seems conclusive to them.

Of course, there are detractors. For instance, Professors Allen and Pardo seem to deny the first premise, but their main criticism of likelihoods or Bayesian reasoning boils down to the observation that probabilities are not self-defining—they must be estimated on the basis of assumptions and knowledge about the world.²²⁰ This is correct—nature may not be what it seems, ambiguities often prevent agreement on the values of the probabilities, and one can overlook plausible alternative hypotheses in formulating the problem. But these reminders do not refute the claim that evidence is probative to the extent that it alters our best estimate of the prior probability of a proposition.

The second premise—that the order of evidence does not alter its impact on a prior probability (and hence its probative value)—seems true in theory, but perhaps order affects the assessments of real jurors. The first thing they hear may color what follows, or the last may remain more vivid at the time of jury deliberations. If so, the perceived probative value of the trawl in a trawl-then-investigate case could differ from the perceived value in an investigate-then-trawl case even though the conjoined and commuted evidence is identical. But it is not apparent if there is any consistent order effect in practice or how it would translate into any prejudice against defendants. In any event, this psychological theory would not rescue the frequentist argument that a trawl in the former case is less probative than a single-comparison DNA match. That too is an argument about the logical power of evidence without regard to its psychological impact on jurors.

The third premise is more complicated. It has been argued that the very idea of a prior probability is incompatible with legal proof because the probability before hearing any evidence in a case is undefined²²¹ or, according to the Anglo-American presumption of innocence, must be set to zero.²²² Probabilists and philosophers have struggled for centuries over the assignment of partial beliefs or

219. *E.g.*, Berry, *supra* note 215, at 253.

220. Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. LEG. STUD. 107, 112 (2007).

221. Lea Brilmayer & Lewis Kornhauser, *Quantitative Methods and Legal Decisions*, 46 U. CHI. L. REV. 116, 143–44 (1978).

222. *State v. Skipper*, 637 A.2d 1101, 1108 (Conn. 1994).

probabilities in cases of complete ignorance.²²³ Bayesians commonly avoid the issue by asserting that there is always *some* other evidence in a case that can serve as the starting point for the prior probability distribution.²²⁴ At a minimum, the jury will have some indication of the defendant's connection (or lack of it) to the place where the crime occurred and of the defendant's physical and mental capacity to have committed the alleged offense.²²⁵ Even starting with prior odds as low as one to a million, the DNA evidence can produce appreciable posterior probabilities, as we saw in the naked-trawl example.

Moreover, in the context of criminal trials, a normative basis for assigning prior probabilities in advance of any evidence is available.²²⁶ Arguably, the presumption of innocence means that, *ab initio*, everyone in the jurisdiction, the country, or even the world, should be regarded as having the same prior probability.²²⁷ For a suspect population of size N , this interpretation is tantamount to using a uniform prior probability of $1/N$.²²⁸

Some courts have expressed reservations or antipathy toward the explicit use of Bayes' rule in the trial process, and proposals to tutor jurors in how to adjust prior odds have been the subject of an extended debate among legal commentators.²²⁹ The California Court

223. See, e.g., Max Black, *Probability*, in THE ENCYCLOPEDIA OF PHILOSOPHY 464, 474 (Paul Edwards ed., 1967) (discussing "the principle of indifference" associated with Laplace and Bernoulli).

224. Cf. *id.* ("It is hard, of course, to imagine a case in which the reasoner is *wholly* ignorant of evidence favoring either P or not- P .").

225. Cf. David H. Kaye, *The Laws of Probability and the Law of the Land*, 47 U. CHI. L. REV. 34, 44-45 n.37 (1979) ("[T]here will almost always be some information on which to base a subjective probability.").

226. There is no legal principle that forces zero to be the prior probability before evidence is heard. If zero were required, Bayes' rule would imply that no conviction is possible (since multiplying by zero always gives zero for a posterior probability). But treating any empirical proposition as having a probability of zero or one amounts to a dogmatic assertion that no amount of evidence to the contrary can change the probability. LINDLEY, *supra* note 190, at 104. This hardly seems appropriate in a system in which jurors are asked to approach the case with an open mind.

227. See KAYE ET AL., *supra* note 18, § 12.4.3.

228. This assumes that a crime has occurred. The probability of $1/N$ could be reduced to reflect the probability that no one is guilty because no crime has been committed. What the latter probability might be in the absence of all evidence is obscure, but all that matters for present purposes is that the presumption of innocence plausibly can be construed as meaning that, initially, the defendant is exchangeable with everyone else in the population. Under this interpretation of the legal norm, the prior probability distribution is uniformly low, perhaps even lower than $1/N$.

229. See 1 MCCORMICK, *supra* note 96, § 210.

of Appeal in *People v. Nelson*²³⁰ disparaged the Bayesian solution to the trawl problem on this ground. The court wrote that:

Use of a Bayesian formula requires a quantified prior probability and quantifiable new information. . . . Bayesian analysis then utilizes a complicated formula to revise the prior probability on the basis of the new information Bayesian techniques are inherently confusing and would be difficult, if not impossible, to explain to an average jury The end result of a Bayesian analysis is often misleading.²³¹

Admittedly, “whether the benefits of using [Bayes’ rule] solely to educate the jury by displaying the probative force of the evidentiary findings would be worth the costs in terms of time-consumption and possible confusion is a [close] question,”²³² but it is not a question that arises in using Bayes’ rule to decide whether trawling degrades the probative value of a DNA match. The argument is not that the jury must use Bayes’ rule to give a trawl its proper weight. Instead it is that the jury, whether or not it is told that the defendant was identified through a trawl, can use the random-match probability just as it would in a confirmation case.

This leaves but one possible flaw in the argument about the probative value of a trawl. The third premise implicit in our use of Bayes’ rule is the postulate of equal priors as between trawl and confirmation cases. This works for an idealized juror who reasons as Bayes’ rule prescribes. But would real jurors be prone to overvalue the DNA match by, in effect, using too high a prior probability in a trawl-then-investigate case? The FBI’s DNA Advisory Board apparently thought so, at least in the context of a naked-trawl case.²³³ But this impression of how jurors reason is debatable. Without a rather subtle empirical study of real jurors, it is difficult to say how

230. 48 Cal. Rptr. 3d 399 (Ct. App. 3 Dist. 2006), *aff’d*, 185 P.3d 49 (Cal. 2008).

231. *Id.* at 416. Strangely, the court cited my revision of MCCORMICK ON EVIDENCE § 211 at 335 (John Strong ed., 5th ed. 1999), which only condemned a dubious application of Bayes’ theorem in parentage testing cases. With regard to correct applications of the theorem in that context, I wrote that “[t]here is . . . a strong argument for using a Bayesian approach to help the jury evaluate the evidence.” *Id.*

232. 1 MCCORMICK, *supra* note 96, § 210 at 332.

233. The Board wrote that:

[W]ithout the Bayesian framework, the Balding and Donnelly (1996) formulation is easily misinterpreted in a fashion unfavorable to the suspect. Stockmarr’s (1999) formulation, which is a more formal exposition of what originally appeared in the NRC II Report (1996), communicates a value of a database search far better, and it is always conservative.

DNA Advisory Board, *supra* note 22.

jurors handle naked DNA evidence cases. In any event, as I argue below, the defense should be permitted to argue that the prior probability is so small that the DNA match does not establish guilt.²³⁴ In these circumstances, the normal rules of evidence should allow jurors to learn of a match and the random-match probability instead of substituting (or adding) an inflated match probability.

Thus, all the premises of the Bayesian analysis of trawling emerge intact. The fact that the defendant is identified by a trawl strengthens rather than weakens the hypothesis of guilt. The random-match probability therefore retains its value as an indication of probative value, and courts like *Jenkins*, *Johnson*, and *Nelson* are justified in admitting random-match probabilities even in trawl cases. But this is only the prosecution side of the equation. The final Part of this Article turns to the defense side. It discusses how the defense can place the random-match probability in proper perspective.

III. BAYES IS BETTER?

Some statisticians have proposed that juries should be advised to start with a uniform prior distribution and adjust this prior iteratively for each new item of evidence.²³⁵ This is quite different than the preceding analysis which used Bayes' rule solely as a heuristic device to analyze whether a trawl is less probative than an isolated match and did not demand that an idealized juror update a personal probability for every new piece of evidence.²³⁶ Nevertheless, as trawls continue to generate suspects and prosecutions, defendants may be tempted to pursue this approach. Several courts have suggested that this defense tactic should not be permitted. As we just saw, *Nelson* contains such dicta, and as we soon shall see, the English Court of Appeal has been hostile to the courtroom use of Bayes' rule. Despite these judicial misgivings, Section A argues that defendants should be allowed to advance an explicitly Bayesian treatment if they choose. It also describes how the prosecution can respond. Section B argues against the less coherent approach of presenting both the adjusted and the unadjusted probability, as suggested in *Jenkins* and *Nelson*. It contends that multiplying by the database size n produces too large

234. See *infra* Part III.A.

235. E.g., Berry, *supra* note 215, at 254 (proposing one possible initial distribution).

236. Of course, one can modify the details of this proposal to allow the jurors to think about all the nongenetic evidence as a whole, then apply Bayes' rule. Indeed, that is the manner in which Finkelstein & Fairley, *supra* note 34, first provocatively proposed incorporating Bayes' rule into the trial process.

an adjustment and sketches a frequentist alternative that is more appropriate.

A. Bayes for the Defense

Although I have deployed Bayes' rule to demonstrate that a trawl match is at least as probative as a confirmation match, thereby supporting the prosecution's use of the random-match probability in trawl cases, the rule could be employed by a defendant to show that the trawl does not give the jury sufficient evidence to convict. Defense counsel pursued the strategy of using Bayes' theorem with a uniform prior probability distribution in two unusual trials stemming from a rape north of London. In its first opinion in *R. v. Adams*,²³⁷ the Court of Appeal related that "a Miss Marley . . . was walking home after an evening out on April 6, 1991. Her attacker was a stranger. He approached and asked her the time. She saw his face for a matter of seconds before looking at her watch. He raped her from behind."²³⁸ She described her assailant to the police as "a white, clean shaven, man with a local accent aged 20 to 25."²³⁹ Time passed. "In October 1993 she attended an identification parade but did not pick out [Adams] or anyone else."²⁴⁰ In fact, she later said that Adams did *not* look like the man who had attacked her.²⁴¹

The prosecution's case rested entirely on a match between a DNA sample from Denis Adams and semen from a vaginal swab taken soon after the attack. The Crown's expert testified that "the chance of a randomly chosen unrelated man matching the DNA profile was one in 297 m. rounded down in the interests of 'conservatism' to 200,000,000."²⁴² Adams maintained that the match was coincidental. He had a brother, and an expert testified that the chance of "a full brother having the same nine DNA bands is one in 220."²⁴³ He testified that he could not have been the rapist because he had spent the night with his girlfriend, who corroborated his testimony.²⁴⁴ Peter Donnelly, Professor of Statistical Science at Oxford University, testified for the defense. He expressed concerns about sampling error in estimates of allele frequencies and other

237. [1996] 2 Crim. App. 467 (CA (Crim. Div)).

238. *Id.* at 468.

239. *Id.*

240. *Id.*

241. *Id.*

242. *Id.*

243. *Id.*

244. *Id.*

matters, and suggested that “the right answer might be in the range of 2 m. rather than 200,000,000”²⁴⁵

In addition, Donnelly described to the jurors how they might apply Bayes’ rule.²⁴⁶ Rather than begin with a prior probability reflecting all the nongenetic evidence—the circumstances of the crime, the fact that Adams did not match the victim’s description of the assailant, and his alibi—Donnelly started with a prior probability based on the general circumstances alone. He noted that in “Hemel Hempstead there are just over 150,000 . . . men between the age of 18 and 60.”²⁴⁷ Assuming as “an illustration,” that it was 75% probable that the rapist came from this region, Donnelly arrived at 200,000 for the suspect-population size. In this manner, he arrived at the prior “odds on a particular local man . . . being the true rapist [of] 200,000 to one”²⁴⁸

Starting from this uniform prior, Donnelly used examples of likelihood ratios for each item of evidence—the victim’s inability to identify Adams, Adam’s testimony, the girlfriend’s testimony, and the DNA match—in successive iterations of Bayes’ rule to conclude that the final posterior odds would be 55:1 on guilt if the random-match probability were 1/200,000,000, but only 5.5:1 if that probability were 1/20,000,000, and only 0.55:1 (“he is almost twice as likely to be innocent than guilty”) if the match probability were 1/2,000,000.²⁴⁹ As Professor Donnelly wrote years later, “unlike in many DNA cases, it really did matter whether the number was 1 in 2 or 20 or 200 million.”²⁵⁰

Despite this presentation, the jury convicted Adams. Adams appealed. He argued, first, that “the DNA evidence upon which the Crown had relied was incapable on its own of establishing guilt.”²⁵¹ The Court of Appeal (like its counterparts in America)²⁵² saw nothing

245. *Id.* at 470.

246. Professor Donnelly distanced himself from this presentation, writing that “I should be clear that it was neither my suggestion nor my choice to explain Bayes’s [sic] Theorem to the jury in the Adams case, and I remain unconvinced that it is a practicable way forward. In Adams, it was the defence barrister’s choice, presumably because he felt it would increase the possibility of a favourable verdict for his client.” Peter Donnelly, *Appealing Statistics*, 2 SIGNIFICANCE 46, 48 (2005).

247. *Adams*, 2 Crim. App. at 472.

248. *Id.* at 473–74.

249. *Id.* at 477.

250. Donnelly, *supra* note 246, at 47.

251. *R. v. Adams* [1998] 2 Crim. App. 377, 378 (CA (Crim Div)).

252. See *Roberson v. State*, 16 S.W.3d 156, 170 (Tex. Crim. App. 2000) (observing that “the perils of eyewitness identification testimony far exceed those presented by DNA expert testimony” and affirming that a verdict can be based on DNA alone (quoting

inherently unjust or improper in basing a conviction on a naked DNA match.²⁵³ But Adams argued, second, that the trial judge, in summing up the evidence for the jury, failed to recapitulate Donnelly's testimony clearly and accurately and neglected to instruct the jury in how to combine the genetic and nongenetic evidence if it did not wish to apply Bayes' theorem.²⁵⁴ Now the Court of Appeal agreed.²⁵⁵ It quashed that conviction and ordered a retrial. Furthermore, although there had been no objection to the reliance on Bayes' rule at the trial and no argument on appeal that the testimony was inadmissible, the Court of Appeal *sua sponte* opined that "to introduce Bayes Theorem, or any similar method, into a criminal trial plunges the jury into inappropriate and unnecessary realms of theory and complexity deflecting them from their proper task."²⁵⁶

So Adams faced a second trial. The defense was unwilling to drop the Bayesian analysis. The prosecution's statisticians from the Forensic Science Service agreed that "in principle Bayes's [sic] Theorem was the logically correct way to combine . . . evidence."²⁵⁷ Despite the ominous dicta from the appellate court, the trial judge agreed to allow a Bayesian presentation. Indeed, he "asked the experts on both sides to get together and prepare a questionnaire which would help the jury to implement Bayes's [sic] Theorem, should they choose to do so."²⁵⁸ In due course,

[t]he questionnaires were produced, and there were boxes where [the jurors] could enter their numerical assessments, with a formula explaining how to combine them. The jury were told that this was in the experts' view the right way to do the

People v. Rush, 630 N.Y.S.2d 631, 634 (N.Y. Sup. Ct. 1995), *aff'd*, 672 N.Y.S.2d 362 (App. Div. 1998)); *People v. Rush*, 672 N.Y.S.2d 362, 363 (App. Div. 1998) (upholding conviction based only on DNA evidence, even given that complainant misidentified defendant at trial, and rejecting argument that DNA is not "infallible" and thus cannot stand alone because "[v]irtually no evidence is absolutely conclusive").

253. But see Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CAL. L. REV. 721, 741 n.87 (2007) (referring to an unreported English Court of Appeal case quashing a conviction where the random match probability of profile was about one in four million and the accused, according to Mike Redmayne, *Appeals to Reason*, 65 MOD. L. REV. 19 (2002), had no connection to the area).

254. *R. v. Adams* [1998] 2 Crim. App. 377, 378 (CA (Crim Div)).

255. *Id.* at 385.

256. *Id.* at 384 (quoting *Adams*, 2 Crim. App. at 482).

257. Donnelly, *supra* note 246, at 47.

258. *Id.*

reasoning, but that they were the jury and it was entirely up to them²⁵⁹

Furthermore, the court supplied basic electronic calculators to the jurors, and Donnelly “walked the jury through a numerical example—the barrister would suggest token numbers in answer to the questions, and the jury and I entered them in the calculators. . . . They seemed to have no difficulty in following this”²⁶⁰ The judge instructed the jury to take the questionnaires to the jury room as well as a blank one in case they wished to enter a prior probability and likelihoods “to fill in your collective view if you want to.”²⁶¹ But he did not encourage them to deliberate in this fashion. To the contrary, he informed them that:

There is absolutely no compulsion on you to use it at all. . . . It was suggested by [defendant’s counsel] that you might think it only fair to this defendant for at least one of you to do it. I hope he will forgive me if I discourage that and for this reason: Your duty, when it comes first thing tomorrow morning, is to retire, consider your verdict amongst yourselves, all of you together and not with one huddled in a corner with his calculator.²⁶²

The jury returned a verdict of guilty, and Adams appealed. His argument that the trial judge was too disparaging of the questionnaires failed dismally. Again, and in the same words as before, the Court of Appeal condemned the elaborate exercise in Bayesian reasoning as confusing and distracting, plunging “the jury into inappropriate and unnecessary realms of theory and complexity deflecting them from their proper task.”²⁶³

259. *Id.* at 48.

260. *Id.*

261. *Adams*, 2 Crim. App. at 382.

262. *Id.*

263. *Id.* at 384. Arguably, the court left room for a Bayesian presentation in extraordinary circumstances. It wrote that:

We do not consider that [jurors] will be assisted in their task by reference to a very complex approach which they are unlikely to understand fully and even more unlikely to apply accurately, which we judge to be likely to confuse them and distract them from their consideration of the real questions on which they should seek to reach a unanimous conclusion. We are very clearly of opinion that in cases such as this, lacking special features absent here, expert evidence should not be admitted to induce juries to attach mathematical values to probabilities arising from non-scientific evidence adduced at the trial.

The court's disapproval of a Bayesian presentation by the defense is too sweeping. The jury's task is to give the DNA evidence the weight it logically deserves. Although the *Adams* case, with its calculators and sequential processing of all the evidence, understandably struck the Court of Appeal as problematic, "realms of theory and complexity" need not accompany an illustration of how a likelihood ratio alters a prior probability.²⁶⁴ In a naked-trawl case, for example, it would not be unreasonable for the defendant to argue that the jury should begin by treating everyone in the general population as equally likely to have left the DNA.²⁶⁵ In our naked-trawl modification of the *Carder* case, the jury could be informed of the following: that if one starts by assuming that the defendant is no more likely than any other man in San Diego County to be the source of the semen stain, then the prior probability is about one in 1,400,000; that if the source is unrelated to the defendant, the match to the semen stain is very unlikely (1/3,000,000); and that the match raises the probability to only about 68% (or even less if relatives are considered).

The prosecution might not like this calculation. It might regard as ridiculous the assumption that there were 700,000 men to consider as possible murderers. Indeed, even if this starting point does capture the normative sense of the presumption of innocence, the defense might be foolish to use it. The number applies when the jury knows only that a man sexually assaulted someone in San Diego. As soon as the prosecution proves that a man with certain physical features took Alicia from her bedroom and sexually assaulted her, the jury knows much more. It will have some idea of the man's age, race, and physical abilities. This will exclude many of the 1.4 million men living in San Diego country from the initial pool of men who are exchangeable with Carder. If we limit the pool to the most common race in San Diego County, whites, and to men between the ages of 18 and 65, the population size drops to approximately 450,000, and the probability rises to 87%. Even if Alicia's description of the age were this vague, this population seems too large since some men would have been in the hospital or have been too feeble to have carried her away; many would have airtight alibis; many would be psychologically incapable of the alleged conduct.

Id. at 385. *Adams* was not identified by trawling the British database, but whether a naked trawl would constitute a "special feature" that would justify instructions in mathematics is obscure. *Id.*

264. *Id.* at 384.

265. See *supra* Part II.D.1.

Alternatively, rather than argue that the defendant's selection of a prior probability is unrealistic, the prosecution also could regard many of these men as part of the suspect population but point to admissible facts about Carder that distinguish him from the mass of male humanity. Did he live or work near Alicia's home? Was he in good health with the physical capacity to enter through the window and carry the little girl off? These considerations and others like them will shift additional probability mass to Carder. If Carder spoke to the police or if he takes the witness stand, there would be other reasons to challenge the defense's number. Did he have an airtight alibi, for example, as many of the 1,400,000 men surely did? But if, at the end of the day, the prosecution cannot distinguish Carder from the remainder of the male population that was capable of performing the alleged acts in this naked-trawl case, then he *should* be acquitted.²⁶⁶ In sum, the defendant's effort to say what the match proves might be flawed, but it is within the realm of reason and can stimulate the jury to think more critically about the meaning and limitations of the naked-trawl evidence. It does not seem so complex and esoteric as to be antithetical to the trial process.

B. Other np-like Rules

Using Bayes' rule for the defense by starting with a small enough prior probability in a cold hit case is only one tactic that a defendant could use to try to overcome the prosecution's random-match probability. In *People v. Puckett*, for example, the defendant tried to introduce evidence that the random-match probability pertained to a

266. Moreover, in some circumstances, a gap in the prosecution's case is itself evidence in favor of the defendant's innocence. If one would expect to have the missing evidence when the defendant is guilty and not when he is innocent, it tends to prove innocence. For a Bayesian analysis, see D.H. Kaye, *Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?*, 66 B. U. L. REV. 657 (1986), reprinted in PROBABILITY AND INFERENCE IN THE LAW OF EVIDENCE: THE USES AND LIMITS OF BAYESIANISM 129 (Peter Tillers & Eric Green eds., 1988). For example, if a bank has been robbed and a suspect has been apprehended before he could have hidden or disposed of the currency, the failure of a diligent search to locate the currency in his possession should give us pause. Thus, naked-trawl evidence is weaker when there ought to be more to the case than a DNA match.

In some naked-trawl cases, of course, the prosecution can explain away the gap in its case. Skillful advocates appreciate the need to explain the absence of expected evidence, and the law of evidence should allow for such proof in cases of negative evidence. The case against John Davis mentioned at the start of this Article, for example, was cold as well as naked. Barbara Martz had been murdered over 20 years ago. The absence of a witness or circumstantial evidence besides the DNA that would associate Davis with the death of his neighbor might be expected even on the hypothesis of guilt. See Van Derbeken, *supra* note 6.

database trawl and that the adjusted probability np was not 1 in 1.1 million, but one in three.²⁶⁷ The trial court refused to allow this²⁶⁸ although it is not clear why. Certainly, when the defendant wants to prove that he was identified by a trawl, the argument that revealing the defendant's criminal status would be prejudicial to the defendant makes little sense. Here, I suggest that the 1/3 figure is, strictly speaking, irrelevant, and that if a court were to give the defense the latitude to call experts to testify to adjusted probabilities,²⁶⁹ the np adjustment would need to be modified.

If pure logic were the only consideration, no np -like statistic would be permitted. The figure of $np = 1/3$ in *People v. Puckett*, for instance, is an estimate of the probability that a database of $n = 338,000$ would yield a hit to *someone* (not necessarily Puckett) if it were composed exclusively of individuals who are not the source of the crime-scene DNA (and who are not identical twins of the true source). Unlike the random-match probability of $p = 1/1,100,000$, however, this number is not a likelihood that is of interest to the jury.²⁷⁰ As the California Supreme Court acknowledged in *Nelson*, the legal issue is not whether the database is innocent, but only whether the named defendant is guilty or innocent.²⁷¹ From the Bayesian perspective, it is hard to see how the 1/3 figure is of much benefit to a juror in assessing the hypothesis that Puckett's DNA (as opposed to some unrelated person's DNA) was recovered from the victim.

Nevertheless, defendants can be expected to argue that the innocent-database-match probability is not completely irrelevant. After all, if the database is innocent, then so is the individual who became a suspect because his profile was in it. Although each side in the statistical debate over adjustment claims that only its favorite probability should be admissible, perhaps the introduction of the innocent-database probability on top of the random-match probability can be rationalized as a Solomonic compromise, giving each side half of what it wants and trusting that their further

267. Chris Smith, *DNA's Identity Crisis*, S. F. MAG., Sept. 2008, <http://www.sanfrancisco.com/story/dna-s-identity-crisis>.

268. *Id.*

269. Expert testimony to this effect is (or has been) available to the defense. See, e.g., *id.* (describing the opinions of Professor William Thompson).

270. A "likelihood," it will be recalled from Part II, measures how much a datum supports a hypothesis. Technically, it is proportional to the conditional probability of the evidence given the hypothesis. See, e.g., RICHARD ROYALL, *STATISTICAL INFERENCE: A LIKELIHOOD PARADIGM* 24–28 (1997).

271. *People v. Nelson*, 185 P.3d 49, 66 (Cal. 2008).

statements will clarify the matter for the jury. This let-it-all-in ruling would allow individuals who were charged as a result of a database trawl to introduce testimony about the chance that an innocent database would have within it a profile that matches the crime-scene DNA. Although introducing both the adjusted and the unadjusted number seems more confusing than edifying, it might well appeal to judges who understandably find it hard to decide whether the frequentist perspective is appropriate here.

If defendants are permitted to introduce an adjusted probability—an outcome that certainly is consistent with the (incomplete) analysis in *Jenkins* and *Nelson*—then this adjustment should not be made with the nominal size n of the database. Instead, a smaller, “effective database size,” n' , should be used.²⁷² The frequentist perspective suggests that the jury needs to know how often innocent defendants would be charged if the null hypothesis (that the crime-scene DNA came from someone outside the database) were always true. If individuals who clearly could not have committed the crime will not be charged then there is no risk that they will be falsely convicted. Therefore, to the extent that non-viable suspects are included in the database, np does not provide the risk of a false decision even when the null hypothesis is true. Instead of multiplying by n , we should be multiplying by the number n' , which represents the number of potentially realistic suspects within the database.

With convicted-offender databases, n' can be substantially smaller than n because these databases include samples from offenders who are behind bars and could not realistically be the perpetrators of new crimes in the outside world. They may even include individuals who are dead but whose records have not been purged. When evidence from a decades-old crime is tested, young offenders who could not possibly have committed the crime will be in the database. In *People v. Puckett*, for example, many of the 338,000 individuals in the 2004 database were not even alive when Diana Sylvester was killed over thirty years ago. Had the matching software been designed to skip over their profiles, the outcome would have been no more (or less) probative, but the np rule would have produced a smaller adjusted probability. Nothing useful is accomplished by complicating the database searches so that they

272. I am grateful to John Hartmann for raising this point. E-mail from John Hartmann, Senior Forensic Scientist, Orange County, Cal. Sheriff-Coroner Dep't., to D.H. Kaye, Regents' Professor, Sandra Day O'Connor College of Law (May 23, 2008, 14:41 MST) (on file with the North Carolina Law Review).

exclude impossible suspects before the trawl or look first at the most plausible suspects.²⁷³ If the defendant is allowed to introduce an adjusted random-match probability, an n/p statistic is more appropriate than the np figure.²⁷⁴

Although the relevance of even the n/p figure is doubtful, it is loosely related to yet another approach that places the random-match probability p in perspective. A defendant might try to argue that even though the DNA profile in question is not likely to be shared by many other unrelated people, it could be present in at least a few other individuals in a large population of potential suspects. Instead of asking how many other people in the database could be considered viable suspects, we could ask how many people would fall into this category of matching, unrelated individuals in the vicinity. Perhaps the entire male population of the San Francisco Bay area between certain ages approximates the size of the initial population of potential suspects. If we call this number n^* , then n^*p is an estimate of the number of unrelated individuals within this population who share the DNA profile taken from the victim's body. Census data suggest that there were over two million men between the ages of eighteen and sixty-four years in the Bay area in 2000.²⁷⁵ Since p is about one in 1.1 million, n^*p is approximately two, indicating that Puckett is not the only man in the region who would have the

273. Anticipating that some defendant will introduce the np figure to moderate the prosecution's p , the state might be tempted to adopt a more complicated search strategy. It could arrange the database in order of decreasing suspicion. The least suspicious people in the database would be those who were incarcerated, not yet born, or very young at the time that the crime was committed. They would be at the bottom of the list. Those living in the locale of the crime and being of the same sex, age, or race of the perpetrator (when such information is available) would be at the top of the list. Instead of searching the entire list, the state could confine (at least initially) the trawl to the most likely candidates. Rather than trawl all 338,000 profiles in *Puckett*, for instance, California might have trawled only (let us say) 11,000 profiles and stopped if it obtained a match in that group of prime suspects. In this situation, np would have been $11/1100 = 1/100$. Or, it could have gone down the whole, ordered list until it obtained a cold hit, then stopped. If the hit occurred at profile number 110, for instance, it might seem that the adjusted p would be $110/1,100,000 = 1/1000$. However, this method of adjusting p would not be the correct if the stopping rule would have allowed more profiles to be compared had the cold hit not occurred at that point. See DAVID O. SIEGMUND, *SEQUENTIAL ANALYSIS: TESTS AND CONFIDENCE INTERVALS* 37–39 (1985); Thomas D. Cook, *P-Value Adjustment in Sequential Clinical Trials*, 58 *BIOMETRICS* 1005, 1005 (2002) (discussing the effect of stopping rules in clinical trials in medicine).

274. Cf. Storvik & Egeland, *supra* note 22, at 923–24 (describing but not endorsing a substitute for the np rule that incorporates the size of a realistic population of potential suspects).

275. METRO. TRANSP. COMM'N AND ASS'N OF BAY AREA GOVT'S, *BAY AREA CENSUS* (2003), available at <http://www.bayareacensus.ca.gov/bayarea.htm>.

requisite DNA profile. This reasoning is a much simplified version of the Bayesian presentation, outlined in the previous Section. In essence, it postulates a uniform prior probability distribution over most of the male population in the geographic region. As noted there, it remains open to the prosecution to argue that the defendant is much more likely than the other possible members of this suspect population—as whittled down by the DNA evidence—to be the source. But if the prosecution has no evidence or reasonable arguments that can distinguish the defendant from the genetic cohorts, then the jury lacks a posterior probability large enough to warrant a conviction.

CONCLUSION

The statistical literature on database trawls reflects two different perspectives on proof. The frequentist perspective that has motivated recommendations to devalue a match in a database is appropriate when asking how often trawls of an “innocent database” can be expected to incriminate a person who is genetically unrelated to a crime-scene stain. But this is a poor measure of the probative value of a database search in comparison to a confirmation search. The database match is at least as probative as the more limited information that comes from an isolated match. Consequently, the fear of data mining is not a reason to present an inflated random-match probability.

The trawl evidence, standing alone, may *seem* weaker because a naked-trawl case has less evidence *in toto* than a confirmation match. But regardless of whether a trawl comes before, after, or without substantial nongenetic evidence, a complete trawl of a large database produces a large change in the prior odds. As such, it provides probative evidence. Because the change in the probability is slightly greater than that for a nontrawl DNA match to the same individual, the case can be presented, if need be, without mentioning the trawl and without adjusting the random-match probability.

Courts have been reluctant to pursue the reasoning that divides frequentists and Bayesians. They have recognized that the question that is posed determines the statistical answer, but they have not penetrated to the core issue of which question—and hence, which statistic—is relevant. The argument developed here supports the admission of the random-match probability in database-trawl cases. As the *Jenkins* and *Nelson* courts recognized, the issue is indeed one of relevance and legal policy, not genetics and DNA science. However, the statistical literature on multiple-hypothesis testing and

database searches is pertinent to this legal question. Only by sorting out the logical arguments for and against adjustment of the random-match statistic can the trawl issue be put to rest.²⁷⁶

276. In pursuing these arguments, I have concluded that whatever numerical value the random-match probability for a given profile might have, an objection to the random-match probability on the ground that trawling degrades the probative value of the match is not well taken. Of course, the analysis does not address other possible objections to random-match probabilities. For example, I have taken estimates of random-match probabilities in the trillionths, quadrillionths, quintillionths, sextillionths, and septillionths at face value. Such exquisitely small numbers result primarily from the population-genetics models that justify the repeated multiplication of allele frequencies. The models have been subjected to considerable study and found to be reasonable, but whether they can be pushed to these limits is open to debate. This question, and others like it, must be left for another day.

APPENDIX

This appendix provides a more precise and compact mathematical analysis of the effect of a database tawl.²⁷⁷ It uses the notation of Part II, which is summarized in Table A.1:

Table A.1. Notation for
Hypotheses and Data in a Database Trawl Case

Individuals in the Population						
[I_0		I_1		I_2 ... I_{N-1}]
[Defendant		The rest of the population of size N]
Individuals in the Database						
[I_0		I_1		... I_{N-1}]
[Defendant		The rest of the database of size n]
Hypotheses About the Source						
[S_0		S_1		S_2 ... S_{N-1}]
Likelihoods for Hypotheses About Everyone						
[L_0		L_1		L_2 ... L_{N-1}]
[$\Pr(E S_0)$		$\Pr(E S_1)$		$\Pr(E S_2)$... $\Pr(E S_{N-1})$]
Prior Probabilities						
[π_0		π_1		π_2 ... π_{N-1}]
[$\Pr(S_0)$		$\Pr(S_1)$		$\Pr(S_2)$... $\Pr(S_{N-1})$]
Posterior Probabilities						
[π_0'		π_1'		π_2' ... π_{N-1}']
[$\Pr(S_0 E)$		$\Pr(S_1 E)$		$\Pr(S_2 E)$... $\Pr(S_{N-1} E)$]
Events (Facts or Data) and Evidence (E)						
M_j	An individual I_j has the same genotype as the crime-scene sample (a match)					
X_j	An individual I_j has a different genotype as the crime-scene sample (an exclusion)					
E_n	The event that defendant matches and all the other individuals in a database of size n are excluded (M_j & X_1 & ... & X_{n-1})					

277. The exposition in Part II elaborates on Walsh & Buckleton, *supra* note 22, at 464–69.

A. Derivation of Bayes' Rule

The world can be in one of N possible, mutually exclusive states S_j ($j = 0, 1, \dots, N-1$). We make an observation E that has a well-defined probability $\Pr(E|S_j)$ of arising under each state. For example, a card drawn at random from a well shuffled deck of playing cards might be an ace of diamonds. This is one of the fifty-two most finely grained, possible states. The observation E might be that the card is an ace. But we have no data on which ace it is.

Let $\Pr(S_j)$ be the probability of the state prior to making the observation. Let $\Pr(S_j|E)$ be the conditional probability of the state given the observation E . The definition of conditional probability states that

$$\Pr(S_j|E) = \Pr(S_j \& E) / \Pr(E). \quad (\text{A.1})$$

The joint probability $\Pr(S_j \& E) = \Pr(E \& S_j)$. Hence,

$$\Pr(S_j|E) = \Pr(E \& S_j) / \Pr(E). \quad (\text{A.2})$$

From the definition of conditional probability, we also know that

$$\Pr(E|S_j) = \Pr(E \& S_j) / \Pr(S_j).$$

Rearranging terms,

$$\Pr(E \& S_j) = \Pr(S_j) \Pr(E|S_j). \quad (\text{A.3})$$

Substituting (A.3) into (A.2) gives

$$\Pr(S_j|E) = \Pr(S_j) \Pr(E|S_j) / \Pr(E). \quad (\text{A.4})$$

But $E = (E \& S_0)$ or $(E \& S_1)$ or ... or $(E \& S_{N-1})$. Since each conjunction is mutually exclusive,

$$\Pr(E) = \Pr(E \& S_0) + \Pr(E \& S_1) + \dots + \Pr(E \& S_{N-1}).$$

Now since $\Pr(E \& S_j) = \Pr(S_j) \Pr(E|S_j)$, we have

$$\Pr(E) = \Pr(S_0) \Pr(E|S_0) + \Pr(S_1) \Pr(E|S_1) + \dots + \Pr(S_{N-1}) \Pr(E|S_{N-1}). \quad (\text{A.5})$$

The conditional probability of the observation E given the state of nature S_j can be called a “likelihood” $L_E(S_j)$, or L_j . The prior probabilities $\Pr(S_j)$ can be abbreviated as π_j . Then

$$\Pr(E) = \sum_0^{N-1} \pi_j L_j. \quad (\text{A.6})$$

Substituting (A.6) into (A.4) and writing $\Pr(S_0) \Pr(E|S_0)$ as $\pi_0 L_0$, we conclude that

$$\Pr(S_0 | E) = \frac{\pi_0 L_0}{\sum_0^{N-1} \pi_j L_j}. \quad (\text{A.7})$$

Restating (A.7) in words gives Equation (2) of Part II.

The other formulas for Bayes’ rule in Part II come from expanding the denominator, $\Pr(E)$ in (A.6) as follows:

$$\Pr(E) = \sum_0^{N-1} \pi_j L_j = \pi_0 L_0 + \sum_1^{N-1} \pi_j L_j = \pi_0 L_0 + \sum_1^{n-1} \pi_j L_j + \sum_n^{N-1} \pi_j L_j. \quad (\text{A.8})$$

B. *The Impact of Trawl Evidence*

Part II described four situations:

- A pure confirmation case: nongenetic information points to the defendant; he is the only one tested, and he matches.
- An investigate-then-trawl case: the nongenetic information is followed by a confirming database trawl that points only to the defendant (everyone else in the database is excluded).
- A trawl-then-investigate-successfully case: the trawl points only to the defendant and is followed by an investigation that uncovers the same confirming nongenetic information—one type of cold-hit case.
- A trawl-then-investigate-unsuccessfully case: the trawl points only to the defendant and is followed by an investigation that uncovers no confirming nongenetic information—another type of cold-hit case.

Part II stated that the second and third cases produce the strongest body of evidence. The first case ranks next, and the last case, which has the least total evidence, comes in last. In every case, however, the trawl is equally significant. The remainder of this Appendix proves these assertions and shows how to obtain the numbers for the examples in Part II.

1. The Pure Confirmation Case: $n = 1$

Bayes' rule is easily applied to the simple confirmation case. According to (A.7) and (A.8),

$$\pi'_0 = \Pr(S_0 | E_1) = \frac{\pi_0 L_0}{\pi_0 L_0 + \sum_{j=1}^{N-1} \pi_j L_j}. \quad (\text{B.1})$$

The DNA evidence is that one person, I_0 , has been tested, and he matches: $E_1 = M_0$. If I_0 is the source (S_0), then his genotype must match the genotype in the crime-scene sample:

$$L_0 = \Pr(M_0 | S_0) = 1. \quad (\text{B.2})$$

The likelihood for every other hypothesis S_j ($j > 0$) is the chance that I_0 would match if the untested man I_j ($j > 0$) were the source. For every unrelated man, this probability is the random-match probability p . The prior probability for each S_j could be different, but

$$\sum_{j=1}^{N-1} \pi_j L_j = \sum_{j=1}^{N-1} \pi_j p = p \sum_{j=1}^{N-1} \pi_j = p(1 - \pi_0). \quad (\text{B.3})$$

Substituting (B.2) and (B.3) into (B.1) yields

$$\pi'_0 = \Pr(S_0 | E_1) = \frac{\pi_0}{\pi_0 + p(1 - \pi_0)}. \quad (\text{B.4})$$

Example 1. Let $\pi_0 = 1/1001$. Substituting this value into (B.4) and simplifying the fraction yields $\pi'_0 = 1 / (1 + 1000p)$. Let the random-match probability $p = 1/3,000,000$. Then

$$\pi'_0 = 1 / (1 + 1/3000) = 3000/3001 = 0.9997,$$

as stated in Part II.

2. The Trawl-as-confirmation Case: $n > 1$

Since only the defendant matches, the DNA evidence is no longer $E_1 = M_0$. Now it is $E_n = M_0 \& X_1 \& \dots \& X_n$ (I_0 matches; everyone else in the database is excluded). This changes the likelihoods in (B.1). Initially, consider a database consisting of only two individuals, I_0 , and an unrelated individual I_1 . The full evidence is $E_2 = M_0 \& X_1$. The likelihood for the defendant, I_0 , is

$$L_0 = \Pr(E_2|S_0) = \Pr[(M_0 \& X_1)|S_0]. \quad (\text{B.5})$$

This quantity can be expressed in terms of the random-match probability. If I_1 is the source, he is certain to possess the matching genotype, but I_1 is not certain to be excluded. I_1 could match by coincidence. The chance of that happening is the random-match probability p . So the chance that I_1 is excluded when I_0 is the source is $1 - p$. This is our likelihood:

$$L_0 = 1 - p. \quad (\text{B.6})$$

We also need the likelihoods L_1 through L_{N-1} . L_1 is easy. If I_1 were the source, he could not have been excluded. Hence, his likelihood is

$$L_1 = 0. \quad (\text{B.7})$$

Moving outside the database to any untested person, I_j ($j > n$), we know that the only way that I_0 can match if I_j is the source is by coincidence, an event that has probability p . And, the chance of excluding I_1 is $1 - p$. Notice also that whether I_1 matches is unrelated to whether I_0 does (when the source is some untested person I_j). The final result is that the likelihood for I_j is

$$L_j = \Pr[(M_0 \& X_1)|S_j] = \Pr(M_0|S_j) \Pr(X_1|S_j) = p(1 - p), \text{ where } 0 < j < N. \quad (\text{B.8})$$

Substituting (B.6), (B.7), and (B.8) into (B.1),

$$\pi'_0 = \Pr(S_0 | E_2) = \frac{\pi_0}{\pi_0 + p \sum_{j=2}^{N-1} \pi_j}. \quad (\text{B.9})$$

The only difference between (B.9) and (B.4) is that in the two-person trawl case, the coefficient of p in the denominator has one less prior in the sum. This reflects the fact that the other member of the database was excluded. A smaller denominator means a larger fraction. The effect of the larger database trawl has been to provide additional information that increases the posterior probability.

The result is easily generalized. If there are n individuals in the database and the trawl excludes all of them except for the defendant I_0 , then the expression for the posterior probability becomes

$$\pi'_0 = \Pr(S_0 | E_n) = \frac{\pi_0}{\pi_0 + p \sum_n^{N-1} \pi_j}. \quad (\text{B.10})$$

To compare (B.10) to (B.4) more transparently, we can use the relationship

$$1 - \pi_0 = \sum_1^{N-1} \pi_j = \sum_1^{n-1} \pi_j + \sum_n^{N-1} \pi_j. \quad (\text{B.11})$$

This equation states that the prior probability not assigned to the defendant ($1 - \pi_0$) is divided into two parts. The second part, which applies to the untested group, is some fraction θ of $1 - \pi_0$:

$$\sum_n^{N-1} \pi_j = \theta(1 - \pi_0). \quad (\text{B.12})$$

The remainder, $(1 - \theta)(1 - \pi_0)$, is the prior probability assigned to the database as a group.

Substituting (B.12) into (B.10), we conclude that

$$\pi'_0 = \Pr(S_0 | E_n) = \frac{\pi_0}{\pi_0 + \theta p(1 - \pi_0)}. \quad (\text{B.13})$$

The only difference from (B.4) is the presence of the fraction θ in the denominator. That $\theta < 1$ reflects the fact that as the database expands and more individuals are excluded as possible sources, more terms disappear from the summation in the denominator, increasing the value of the posterior probability. Trawling itself yields stronger, not weaker evidence.

Example 2. Having identified I_0 by nongenetic evidence, a database of 100,000 people, including I_0 , is searched. Again, let the prior probability for the hypothesis S_0 be $\pi_0 = 1/1001$. Let the fraction

of the remaining prior probability of $1 - \pi_0 = 1000/1001$ that falls on the men outside the database be $\theta = 5/6$. Then, by (B.13), the posterior will be $\Pr(S_0|E_{100,000}) = 18000/18005$. This is slightly larger than $\Pr(S_0|E_1) = 3000/3001 = 18000/18006$ obtained with (B.4). The confirmation-trawl evidence is marginally stronger than the simple confirmation-match evidence.

3. The Trawl-then-confirm-successfully Case

Assume that *after* the trawl points to I_0 , investigators uncover the same clues that would have led them to I_0 in the simple confirmation case, and he is on trial. The order in which jurors process new information according to Bayes' rule (which reduces to (B.13)), makes no difference. Consequently, equation (B.13) applies to the trawl-then-investigate case, just as it does to the investigate-then-trawl case. Again, a successful trawl match is more probative than a successful single-suspect confirmation match.²⁷⁸

To prove explicitly that the same bundle of evidence has the same probative value regardless of whether the trawl comes first or last, we start with a prior probability distribution $\rho_j = \Pr(S_j)$. This distribution applies before considering both the findings F of a conventional investigation and the result T of the trawl. In the investigate-first case, the nongenetic evidence F modifies the prior probability ρ_j that any individual I_j is the source according to Bayes' rule (A.7) with likelihoods $f_j = \Pr(F|S_j)$:

$$\rho'_j = \Pr(S_j | F) = \frac{f_j \rho_j}{\sum_{i=0}^{N-1} f_i \rho_i}. \quad (\text{B.14})$$

Now we trawl and learn that only the defendant matches. To update these probabilities ρ'_j we apply Bayes' rule a second time with the likelihoods $L_j = \Pr(T|S_j)$ for this DNA evidence to obtain:

²⁷⁸ In practice, the order in which investigators marshal evidence can matter. For example, it is conceivable that if an initial trawl would lead the police to the correct suspect sooner, it could allow a more targeted investigation resulting in more powerful nongenetic evidence against the defendant. *See supra* note 172. However, the evidentiary value of the trawl itself is the same whether it comes first or last, and when the nongenetic evidence is the same (as in the analysis here), the order has no effect on the value of the total evidence.

$$\rho_j'' = \Pr(S_j \mid F \& T) = \frac{L_j \rho_j'}{\sum_0^{N-1} L_i \rho_i'} . \quad (\text{B.15})$$

Substituting (B.14) into (B.15) and simplifying yields

$$\rho_j'' = \frac{L_j f_j \rho_j}{\sum_0^{N-1} L_i f_i \rho_i} . \quad (\text{B.16})$$

In the trawl-first, investigate-second case we begin with the prior distribution ρ and update first with the trawl data T to obtain:

$$\rho_j' = \Pr(S_j \mid T) = \frac{L_j \rho_j}{\sum_0^{N-1} L_i \rho_i} . \quad (\text{B.17})$$

Then we update these probabilities with the nongenetic findings F :

$$\rho_j'' = \Pr(S_j \mid T \& F) = \frac{f_j \rho_j'}{\sum_0^{N-1} f_i \rho_i'} . \quad (\text{B.18})$$

Substituting (B.17) into (B.18) and simplifying yields:

$$\rho_j'' = \frac{f_j L_j \rho_j}{\sum_0^{N-1} f_i L_i \rho_i} . \quad (\text{B.19})$$

Equation (B.19) is equivalent to (B.16). This equivalence establishes that the order of acquiring the same evidence has no effect on its probative value, at least where the joint likelihood $\Pr(T \& F \mid S_0) = \Pr(T \mid S_0) \times \Pr(F \mid S_0)$. For such conditionally independent evidence, whether the defendant was a suspect before the trawl, or whether he became a suspect only after the trawl, the combination of the same genetic and nongenetic evidence is equally powerful.

4. The Trawl-then-confirm-unsuccessfully Case

Finally, suppose that the police were unable to locate substantial nongenetic evidence against I_0 . This is the naked-trawl case. If the

prosecution can explain away the lack of other evidence in its case, we need only consider the impact of the evidence E_n on the prior probability π_0 and compare it to the impact of E_1 in the simple confirmation case. Because the hypotheses and the DNA findings are unchanged, there are no new likelihoods to compute. The difference in the nongenetic evidence only affects the prior probability. Because there is no nongenetic evidence of guilt to consider in forming the prior probability in the naked-trawl case, π_0 is smaller than it was in Examples 1 and 2. But whatever its value, the denominator of (B.13) is larger than that of (B.4), and π_0' is correspondingly larger.

Example 3. With no other evidence at hand, a database of 100,000 people, including I_0 , is searched. Let the prior probability for the hypothesis S_0 be $\pi_0 = 1/1,000,001$. Let the fraction θ of the remaining prior probability of $1 - \pi_0 = 1,000,000/1,000,001$ that falls on the men outside the database be $5/6$. Then, by (B.13), the posterior will be $\Pr(S_0|E_{100,000}) = 18/23 = 0.78$.

This is much smaller than $\Pr(S_0|E_1) = 3000/3001 = 0.9997$ obtained with (B.4) in Example 1, but that calculation started with a prior that incorporated the nongenetic information that is not present in this case. If the far smaller prior probability of $1/1,000,001$ is used in (B.4), then the posterior is $\Pr(S_0|E_1) = 3/4 = 0.75$. The correct computation in this example thus shows that the confirmation-trawl evidence is slightly stronger than an isolated DNA match.

One might well ask where the prior probability comes from without some nongenetic evidence to consider. Is there no presumption of innocence? The presumption can be interpreted as follows: the mere fact the defendant has been charged is not evidence. At the outset, a jury should be willing to treat anyone and everyone alike by attaching the same prior probability to everyone in the relevant population. For N people in the population, the uniform prior is $1/N$, and (B.4) simplifies to

$$\pi_0' = \Pr(S_0 | E_1) = \frac{1}{1 + (N-1)p}. \quad (\text{B.4U})$$

The corresponding version of (B.13) is

$$\pi_0' = \Pr(S_0 | E_1) = \frac{1}{1 + \theta(N-1)p}. \quad (\text{B.13U})$$

Because $0 < \theta < 1$, the DNA trawl evidence leads to a higher posterior probability than the DNA evidence in the confirmation case (subtracting out the nongenetic evidence).