SAFETY IN NUMBERS?
DECIDING WHEN DNA ALONE IS ENOUGH TO CONVICT

ANDREA ROTH*

Fueled by police reliance on offender databases and advances in crime scene recovery, a new type of prosecution has emerged in which the government's case turns on a match statistic explaining the significance of a "cold hit" between the defendant's DNA profile and the crime-scene evidence. Such cases are unique in that the strength of the match depends on evidence that is almost entirely quantifiable. Despite the growing number of these cases, the critical jurisprudential questions they raise about the proper role of probabilistic evidence, and courts' routine misapprehension of match statistics, no framework—including a workable standard of proof—currently exists for determining sufficiency of the evidence in such a case. This Article is the first to interrogate the relationship between "reasonable doubt" and statistical certainty in the context of cold hit DNA matches. Examining the concepts of "actual belief" and "moral certainty" underlying the "reasonable doubt" test, I argue that astronomically high source probabilities, while fallible, are capable of meeting the standard for conviction. Nevertheless, the starkly numerical nature of "pure cold hit" evidence raises unique issues that require courts to apply a quantified threshold for sufficiency purposes. I suggest as a starting point—citing recent juror studies and the need for uniformity and systemic legitimacy—that the threshold should be no less favorable to the defendant than a 99.9% source probability.

INTRODUCTION

In 1972, a young nurse was raped and murdered in her San Francisco home. The initial suspect, Robert Baker, had escaped from an asylum a month before, raped a woman less than a quarter-mile from the nurse's home, stalked a young mother and her child's nanny on the nurse's block,[1] and left a parking ticket in his van with blood drops on it matching the nurse's type.[2] But the case against Baker stalled, and the investigation went cold for over thirty years. In 2004, authorities reopened the case, comparing DNA found inside the victim to the 338,000 DNA profiles in California's offender database. The search yielded one match, or "cold hit," to then-seventy-one-year-old John Puckett, a wheelchair-bound man in Stockton.[3] At Puckett's trial, the primary evidence against him was the DNA profile match, though jurors also heard that he lived in the Bay Area in 1972 and had a 1977 sexual assault conviction. The government's DNA expert reported that the chance that a random person from the population would match the profile (the "random match probability," or RMP) was 1 in 1.1 million. The judge allowed the defense to argue to the jury that "[t]here really is no other evidence against Mr. Puckett,"[4] but refused the defense's request to tell the jury that the chance of finding

---

    [1] The mother and nanny stalked by Baker lived at 1240 Sixth Avenue in San Francisco; the nurse-decedent lived at 1279 Sixth Avenue. Appellant's Opening Brief at 81, People v. Puckett, No. A121368 (Cal. Ct. App. Apr. 18, 2009).

    [2] Chris Smith, *DNA's Identity Crisis*, S.F. MAG., Sept. 2008, http://www.sanfranmag.com/story/dna's-identity-crisis. The blood-stained notice was destroyed by the government at some point before trial. Appellant's Opening Brief, *supra* note 1, at 85.

    [3] Edward Humes, *Guilt by the Numbers: How Fuzzy Is the Math that Makes DNA Evidence Look So Compelling to Jurors?*, CAL. LAW., Apr. 2009, at 20, 20–22.

    [4] Appellant's Opening Brief, *supra* note 1, at 56.

at least one match in California's database was one in three.[5] The jury also never heard—and neither party volunteered or asked permission to state—that the chance that another prime-aged man living in the Bay Area in 1972 shared the matching profile was likely over fifty percent.[6] The jury found Puckett guilty, and his appeal is pending.

Cases like *Puckett* are part of an emerging phenomenon of "pure cold hit" DNA prosecutions in which the entirety of the government's case against the suspect, aside from his prior conviction, is a DNA profile match or a match accompanied only by general evidence, such as age or race, that limits the suspect population only to a small and quantifiable degree.[7] These cases have been made possible by law enforcement's increased reliance on DNA databases to investigate crimes, the newfound ability to develop DNA profiles on old, degraded, and mixed evidence samples, and the modern practice of securing so-called "John Doe" indictments against DNA profiles of yet-unknown suspects.[8] Local prosecutors have already brought several such cases, and some have reached appellate courts.[9] While most have ended in a guilty plea, guilty verdict, or acquittal, some pure cold

---

[5] *See* Humes, *supra* note 3, at 22 ("What the jurors didn't know, though, and what the judge didn't think they needed to know, is that there's another way to run the numbers.").

[6] *See infra* note 214 (calculating posterior odds of Puckett being source as only about thirty-three percent); *see also* David H. Kaye, *Rounding Up the Usual Suspects: A Legal and Logical Analysis of DNA Trawling Cases*, 87 N.C. L. Rev. 425, 490–92 (2009) (discussing probability of another source in Bay Area sharing Puckett's DNA profile).

[7] By defining "pure cold hit" in this way, I mean to isolate those cases in which the government's evidence is almost entirely quantifiable, thus raising the issues addressed in this Article related to the limits of purely probabilistic proof of guilt. For example, while the government in Puckett's case introduced not only the profile match but also proof of Puckett's connection to the Bay Area, this added evidence at most limits the size of the likely suspect population to others with a similar connection. Because the chance of another suspect sharing the profile is still relatively high, the case is qualitatively different from cases with other significant individualized evidence of guilt. *See* discussion *infra* Part III (discussing difficulty of finding guilt beyond reasonable doubt in "cold hit" cases). And while a prior conviction might make the suspect statistically more likely to be the perpetrator, courts may not view that fact alone as enough to create a legally sufficient case against the suspect if there are others who also match the profile. *See infra* notes 210–18 (discussing role of evidence of prior conviction in meeting sufficiency standard generally and in *Puckett*). Others have similarly used the term. *See, e.g.*, Simon A. Cole & Michael Lynch, *The Social and Legal Construction of Suspects*, 2 Ann. Rev. L. & Soc. Sci. 39, 49–50 (2006) (describing innocent attorney Brandon Mayfield as "pure cold hit suspect" in Madrid train bombing case where only evidence against him was database match and other details law enforcement found suspicious, for example, that he converted to Islam and had not left country). Still others have used the related term "naked trawl" case to denote a prosecution based on a DNA cold hit (the result of a "database trawl") and no other individualized evidence of guilt. *See, e.g.*, Kaye, *supra* note 6, at 472, 490.

[8] *See* discussion *infra* Part I.B (discussing recent rise in cold hit prosecutions).

[9] *See infra* notes 50–53 and accompanying text (listing cold hit cases).

hit suspects have been exonerated during the investigatory stage upon discovery that the cold hit was coincidental or erroneous.[10]

Because of their starkly numerical nature, their tantalizing offer of—in Justice Alito's recent words—"virtual certainty"[11] of guilt, and the daunting complexity of the statistics involved, pure cold hit cases invite a new conversation about fundamental issues of criminal procedure and evidence law. For example, while a spirited debate persists as to whether culpability should ever be based on naked statistical evidence, pure cold hit cases pose a somewhat different question: whether objections to probabilistic evidence even apply where source probabilities are so high that they are arguably transformed into assertions of certainty rather than probability. If there is a threshold at which the numbers transcend probability, how should courts determine whether the evidence in a given case is legally sufficient to prove guilt? Should judges set a numerical sufficiency threshold? How does the familiar dispute about the propriety of quantifying standards of proof apply to a case in which almost all of the government's evidence is numerical to begin with?

These jurisprudential questions are critical to the future of criminal justice in a technological age; it is thus no surprise that "the question whether or not DNA evidence on its own is enough to convict an accused" was recently described as "one of the most talked-about points regarding DNA evidence."[12] Nevertheless, for being so "talked-about," the question has generated little litigation or scholarship. While scholars have discussed methods of determining whether a profile is unique in a given suspect population,[13] no one has yet suggested a framework for determining sufficiency, including a workable standard of proof, in a pure cold hit case. And while at least one scholar has questioned in passing whether "the sheer statistical power of the probabilities" will transcend arguments that such cases are solely probabilistic and therefore "intrinsically immoral,"[14] no one has addressed the issue in depth. Meanwhile, most courts and litigants grappling with the question of sufficiency have fallen prey to statistical

---

[10] *See infra* notes 54–60 (listing cases and their dispositions).

[11] Dist. Attorney's Office for the Third Judicial Dist. v. Osborne, 129 S. Ct. 2308, 2327 (2009) (Alito, J., concurring) (internal citation omitted).

[12] ANDREI SEMIKHODSKII, DEALING WITH DNA EVIDENCE: A LEGAL GUIDE 136 (2007).

[13] *See infra* note 84 and accompanying text (discussing popularity among scholars of Bayes' Theorem, one method for determination).

[14] Cole & Lynch, *supra* note 7, at 50–51 (quoting Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1372 (1971)). I discuss Tribe's argument more fully in Part III.B. *See infra* text accompanying notes 153–57.

fallacies. Most commonly, they mistake the RMP, the chance that a person randomly selected from the population will match the profile, for the "source probability," the ultimate probability that the defendant is the source of the DNA,[15] and thus fail to correctly assess the strength of the DNA evidence.

My goal in this Article is two-fold. First, I seek to give preliminary answers to the critical, unexplored jurisprudential questions raised by pure cold hit DNA cases. Second, in the process of grappling with these theoretical issues, I hope to illuminate these dramatic cases for their own sake, offering an overview of existing pure cold hit cases and providing a coherent framework for determining sufficiency of the evidence in such cases. Toward these dual ends, Part I describes the pure cold hit phenomenon and reviews existing cases. Part II sets forth the legal sufficiency standard and details how courts have incorrectly assessed sufficiency in pure cold hit cases thus far by focusing on the RMP rather than the source probability. Part III explores the potential critique that pure cold hit evidence is legally insufficient proof of guilt because of its probabilistic nature. Referencing trial-by-mathematics scholarship, cognitive psychology studies of the so-called "Wells Effect,"[16] and the related concepts of "reasonable doubt" and "moral certainty," I argue that there exists a point at which the numbers are so compelling as to amount to an assertion of certainty rather than probability. When that point is reached, pure cold hit evidence is capable of inspiring an actual belief in the defendant's guilt sufficient to justify conviction. In Part IV, I argue that courts should adopt a numerical threshold that reflects this "actual belief" insight and suggest a threshold of a 99.9% source probability. In Part V, I explain how the framework would operate in actual prosecutions, using existing cases as examples. I conclude by raising issues to be explored in future work about the role of the lay jury in cases involving purported near-certain mathematical evidence of guilt.

---

[15] Scientists typically refer to the conflation of the random match probability (RMP) and source probability as the "fallacy of the transposed conditional." *See infra* Part II for further discussion. The RMP of 1 in 1.1 million in *Puckett* does not signify that there is only a 1 in 1.1 million chance that someone other than Puckett is the source of the DNA. Rather, it means that a person randomly selected from the population has a 1 in 1.1 million chance of matching or, equivalently, that we would expect 1 in every 1.1 million people to match the profile. *See infra* text accompanying note 94 (explaining confusion in use of RMP). In a population of about two million men, we would expect about two men to match. In such a case, the probability that the suspect might not actually be the true source might even be greater than fifty percent, a far cry from 1 in 1.1 million.

[16] The "Wells Effect" refers to jurors' demonstrated tendency to impose liability more readily in cases involving impressionistic evidence such as eyewitness testimony rather than naked statistical evidence. *See infra* notes 157–62 and accompanying text (citing work by Gary Wells on this phenomenon).

I

The Emerging Phenomenon of Pure Cold Hit Cases

In this Part, I explain how DNA typing and database development have evolved to the point that a prosecution based solely on a DNA match is possible. I catalog "pure" or nearly pure cold hit cases currently being prosecuted, explain why the number of these cases will likely grow in the future, and note that, with the advent of such cases, journalists and scholars have begun to see the question whether DNA evidence alone can be sufficient to prove guilt as one of critical importance.

A.   Forensic DNA Testing and Use of DNA Evidence at Trial

Two basic types of DNA-based criminal prosecutions exist: "confirmatory" cases and "cold hit" cases. In a "confirmatory" case, a suspect is first identified through non-DNA evidence such as eyewitness identification or a strong motive and then is later determined through forensic testing to match the DNA profile recovered from the crime scene or sample taken from an alleged victim.[17] In a "cold hit" case, by contrast, police develop a DNA profile from an evidence sample but have no identified suspect. The development of DNA offender databases has allowed law enforcement to try to solve such "cold" cases by running the evidence sample profile through such databases to see if the search yields a matching offender profile—a cold hit. Once an offender in the database is identified through a cold hit, he becomes a suspect. In most cold hit cases, police find additional individualized inculpatory evidence, either direct or circumstantial, against the suspect. In other cases, police find little or no additional evidence against the suspect.

During forensic testing, DNA is amplified and typed at several locations, or loci, along the genetic strand. A person's DNA profile consists of two genetic markers, or alleles, at each locus, representing the two alleles he inherits from each of his two parents at that locus.[18] In the currently dominant iteration of DNA-matching technology, short tandem repeat (STR) testing, each allele is a number, repre-

---

[17] The suspect's DNA is sent to the laboratory for testing, either simultaneously or after the evidence sample is tested, depending on the laboratory's practices and the timing of the suspect's arrest. *Cf.* Erin Murphy, *The Art in the Science of DNA: A Layperson's Guide to the Subjectivity Inherent in Forensic DNA Typing*, 58 Emory L.J. 489, 497–98 (2008) (describing how crime scene DNA samples are taken).

[18] *See* John M. Butler, Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers 23–24 (2d ed. 2005) (explaining genetic principles behind DNA testing).

senting the number of times a four-part sequence—such as ATTG[19]—is repeated at a particular locus.[20] Thus, if a person's alleles at a certain locus are 20 and 22, that person inherited a 20-repeat allele from one parent and a 22-repeat allele from the other. State laboratories and the Federal Bureau of Investigation (FBI) typically test thirteen STR loci. If each of the suspect's twenty-six alleles in his DNA profile matches each of the twenty-six alleles in the evidence sample, the laboratory declares a thirteen-loci "match" between the profiles.[21]

In any case involving a DNA match, no matter how the suspect was initially identified, the match is essentially meaningless without some sense of how unusual it is: "Without the probability assessment, the jury does not know . . . whether the patterns are as common as pictures with two eyes, or as unique as the Mona Lisa."[22] To calculate the frequency of the reported profile in the population, the government's DNA analyst consults a statistical table developed by the FBI based on sample groups of about two hundred people from each of four self-reported racial categories.[23] Using the table, the analyst can estimate the chance of finding each particular allele at each particular locus in the different racial groups.[24] Based on the assumption that the allelic frequencies among the loci are statistically independent, the laboratory multiplies the twenty-six frequencies together to report for each group a "random match probability" (RMP), or the probability that a random person selected from the population will exhibit the twenty-six-allele profile.

The size of the RMP depends on how unusual the alleles in the particular profile are, as well as the quality of the evidence sample. Based on allelic frequency tables, the RMP for a fully developed, pristine thirteen-loci STR profile will likely have a denominator at least in the trillions.[25] But some cases involve only partial matches, due to

---

[19] ATTG is a sequence of four units called nucleotides that make up the genetic strand. Here, the sequence includes A (adenine), T (thymine), and G (guanine). *See id.* at 18–19.

[20] *See* Murphy, *supra* note 17, at 494–96 (describing practice of STR testing); *cf.* Butler, *supra* note 18, at 30 (discussing advantages of STR markers).

[21] *See* Murphy, *supra* note 17, at 494–96 (describing testing).

[22] United States v. Yee, 134 F.R.D. 161, 181 (N.D. Ohio 1991).

[23] The categories are African American, Caucasian, Southeast Hispanic, and Southwest Hispanic. *See* People v. Venegas, 954 P.2d 525, 537 (Cal. 1998) (discussing use of categories). Thus far, no court has upheld an admissibility challenge to DNA match statistics based on race-based limitations of the allelic frequency tables.

[24] For example, the estimated chance of a member of the Afro-Caribbean population having a 23 allele at the locus FGA is, according to the FBI's tables, 0.234, or 23.4%. Chantal J. Frégeau et al., *Population Genetic Characteristics of the STR Loci* D21S11 *and* FGA *in Eight Diverse Human Populations*, 70 Hum. Biology 813, 837 tbl.9 (1998).

[25] Bruce Budowle et al., *Source Attribution of a Forensic DNA Profile*, Forensic Sci. Comm., July 2000, http://www.fbi.gov/hq/lab/fsc/backissu/july2000/source.htm ("The

either the low copy number[26] or degraded state[27] of the DNA in the evidence sample. Moreover, in cases involving mixed samples with more than one contributor, the "combined [p]robability of [i]nclusion"[28]—the probability that a randomly selected person's alleles would be consistent with the many alleles in the sample—will be higher than an RMP for a single-source profile.[29] And should the government ever venture beyond nuclear DNA and rely upon a "pure" *mitochondrial* DNA (mtDNA) sequence match to prosecute someone for a crime, the probability of inclusion would also typically be higher than RMPs for nuclear DNA profiles.[30] Finally, the RMP for persons related to the perpetrator is much higher than an RMP assuming a population of unrelated individuals. For example, the chance that two siblings match at all thirteen STR loci can be as high as 1 in 40,000.[31]

---

average random match probability for unrelated individuals for the 13 STR loci is less than one in a trillion, even in populations with reduced genetic variability . . . .").

[26] In low copy number (LCN) cases—involving a very small amount of recovered DNA—typing a full thirteen-loci profile may be impossible. *See, e.g.*, Nuffield Council on Bioethics, The Forensic Use of Bioinformation: Ethical Issues 19, 23 (2007) [hereinafter Nuffield Report] (discussing problems with LCN analysis). At least one court has ruled LCN testing results inadmissible on the ground that the method has not gained "general acceptance" in the relevant scientific community as required in some jurisdictions—including California—under the *Frye* test. Court Reporter's Notes at 3, People v. Espino, No. NA076620 (Cal. Super. Ct. Mar. 18, 2009) (ruling LCN results inadmissible because of lack of consensus in scientific community). The number of such cases is increasing. *See* Cole & Lynch, *supra* note 7, at 43–44 (discussing potential risks of enhanced ability to use LCN samples).

[27] *See, e.g.*, Humes, *supra* note 3, at 23 (discussing 5.5 loci profile due to degradation in *Puckett*); Solomon Moore, *Damaged DNA Evidence Shrinks Serial Killer Case*, N.Y. Times, May 22, 2009, at A14 (discussing partial matches resulting from damaged DNA sample).

[28] Ray Wickenheiser, *General Guidelines for Categorization and Interpretation of Mixed STR DNA Profiles*, 39 Can. Soc. Forensic Sci. J. 179, 214 (2006).

[29] *See, e.g.*, *id.* at 199–207 (explaining example of mixed sample and finding that correct match can be more accurately deduced from single-source sample); William C. Thompson, *The Potential for Error in Forensic DNA Testing*, GeneWatch, Nov.–Dec. 2008, at 5, 6 (noting that chance of coincidental match will increase if sample is mixed compared with single-source sample).

[30] The power of mtDNA in identifying suspects is limited because it is inherited only through the mother, sequence types tend to cluster geographically, and population databases are very limited. *See* Frederika A. Kaestle et al., *Database Limitations on the Evidentiary Value of Forensic Mitochondrial DNA Evidence*, 43 Am. Crim. L. Rev. 53, 66–67 (2006) (identifying problems with mtDNA database). As a result, mtDNA match statistics are currently much less discriminating than RMPs associated with nuclear DNA typing. *See* Semikhodskii, *supra* note 12, at 68–71 (describing different calculations necessary to interpret mtDNA evidence).

[31] *See* Budowle et al., *supra* note 25 ("[A]mong African Americans, Chinese, and Caucasians, the most common conditional probability for a 13 STR locus profile is expected to occur with a frequency no more than one in 40,000 among full siblings.").

At a criminal trial involving DNA match evidence, the prosecution presents evidence to establish that a crime was committed and that the perpetrator left DNA at the scene under circumstances precluding mere innocent presence.[32] The prosecution often calls a serologist or other technician to the stand to explain how the DNA was recovered and extracted and to establish a chain of custody. In other cases, the government has instead introduced the serologist's hearsay report through another expert, such as the analyst who performed the DNA testing—a practice now presumably unconstitutional.[33] The prosecution then typically calls the DNA analyst, who explains both the Polymerase Chain Reaction (PCR)-STR testing process and how she determined which STR alleles were present in each sample based on electropherograms, the graphs that are created by the testing software and that show the results of the testing.[34] The government might also elicit from the analyst the method by which the government determined the RMP via the allelic frequency tables and product rule.[35] In some cases in which the RMP is sufficiently small, the analyst is permitted to assert that, "with reasonable scientific certainty, a particular individual is the source of an evidentiary sample."[36] The analyst decides, based on the practice of her laboratory, when the RMP is small enough to justify such an assertion of source attribution. The FBI currently sets its source attribution threshold at around 1 in

---

[32] As a practical matter, the nature of the evidence in most pure cold hit cases makes mounting an innocent presence defense at trial untenable. *See, e.g.*, Yun S. Song et al., *Average Probability that a "Cold Hit" in a DNA Database Search Results in an Erroneous Attribution*, 54 J. FORENSIC SCI. 22, 23 (2009) ("In cases of sexual assault, courts have reasoned that the intimate nature of the sample forecloses arguments that it might have been left accidentally or inadvertently.").

[33] *See* Melendez-Diaz v. Massachusetts, 129 S. Ct. 2527, 2531–32 (2009) (holding that drug analyst's hearsay report was inadmissible under Confrontation Clause, but noting that not all chain-of-custody witnesses or equipment technicians need testify). Even before *Melendez-Diaz*, several lower courts had recognized that this practice violated the rule of *Crawford v. Washington*, 541 U.S. 36, 68 (2004), which prohibited the government from introducing hearsay affidavits in lieu of live testimony absent a showing of the live witness's unavailability and a prior opportunity for cross-examination. *See, e.g.*, Roberts v. United States, 916 A.2d 922, 938–39 (D.C. 2007) (holding unconstitutional, but affirming due to lack of plain error, government's introduction of serologist's hearsay report through DNA expert).

[34] *See, e.g.*, Appellant's Opening Brief, *supra* note 1, at 11–13 (describing DNA analyst's testimony about testing process); Official Transcript of Proceedings at 3-79 to 3-80, State v. Derr, No. 08-K-04-000930 (Md. Cir. Ct. June 28, 2006) (testimony of FBI Analyst Jennifer Luttman) [hereinafter Luttman Testimony] (describing DNA-testing process).

[35] *See, e.g.*, Appellant's Opening Brief, *supra* note 1, at 17 (describing analyst's testimony regarding RMP calculation).

[36] Budowle et al., *supra* note 25; *see also* D.J. Balding, *When Can a DNA Profile Be Regarded as Unique?*, 39 SCI. & JUST. 257, 258 (1999) (discussing whether sufficiently small RMP can justify calling defendant's profile unique).

280 billion, or about a thousand times the population of the United States.[37] In cases involving a controversy as to how to express the statistical significance of a cold hit match, the government may also call a statistician or population geneticist, either in its own case or to rebut a defense expert.[38]

Defendants are able to cross-examine the DNA analyst about various aspects of the testing process and (upon request) are typically given a compact disc containing all the electronic data generated by the analysis software before the trial.[39] However, they have only a limited ability to meaningfully challenge the two critical statistical factors determining the probative value of a DNA profile match: the chance of a "false" match due to laboratory or interpretive error and the chance that the observed profile match is a "true" match but nonetheless coincidental. Most DNA laboratories do not generate statistics regarding false positive rates of DNA testing procedures, nor do they conduct the type of frequent, double-blind proficiency testing that would be a first step toward generating accurate false positive rates.[40] And while the defense has access to the allele frequency tables used to generate the RMP for various racial groups,[41] it cannot access the millions of DNA profiles in state or federal offender databases to challenge whether the match might be a coincidence. In the wake of evidence that certain state databases have a surprising number of nearly matching profiles,[42] many scholars—including myself—have recently argued that outside researchers must have limited access to large databases in order to meaningfully assess the accuracy of the

---

[37] *See* Luttman Testimony, *supra* note 34, at 3-79 to 3-80 (testifying that when probability is "exceedingly rare," meaning "smaller than 1 in 280 billion in all four of the population databases that we examine," she "will say the person is the source of the DNA to a reasonable degree of scientific certainty").

[38] *See, e.g.*, Smith, *supra* note 2 (noting that government called genetics and statistics expert Dr. Ranajit Chakraborty in *Puckett* case to rebut defense expert).

[39] *See* William C. Thompson et al., *Evaluating Forensic DNA Evidence, Part 2*, CHAMPION, May 2003, at 24, *available at* http://www.nacdl.org/public.nsf/freeform/championmag?OpenDocument (roll over "Champion Magazine" and follow "Search All Articles" hyperlink to find article) (describing ease of providing laboratory files on compact discs for defense attorney review).

[40] *See* Jonathan J. Koehler, *Fingerprint Error Rates and Proficiency Tests: What They Are and Why They Matter*, 59 HASTINGS L.J. 1077, 1094–98 & 1096 n.68 (2008) (discussing design of blind proficiency testing in DNA laboratories while noting difficulty in conducting enough tests to report accurate laboratory-specific error rates); *see also infra* note 204 (discussing importance of false positive rate as counterpoint to RMP).

[41] *See* Bruce Budowle & Tamyra R. Moretti, *Genotype Profiles for Six Population Groups at the 13 CODIS Short Tandem Repeat Core Loci and Other PCR-Based Loci*, 1 FORENSIC SCI. COMM. (1999), http://www.fbi.gov/hq/lab/fsc/backissu/july1999/budowle.htm (presenting genotype frequency tables for thirteen core loci by population group).

[42] *See infra* note 203 (discussing criticisms of reported RMPs and reports that matches are less rare than claimed).

government's claims that thirteen-loci STR profiles are so rare as to be most certainly unique in the population.[43]

B. The Rise in "Cold Hit" and "Pure Cold Hit" Prosecutions

When national DNA databases were first developed in the 1990s,[44] the number of cold hit prosecutions—those in which the suspect was identified through a database match—was relatively low. In the last decade, however, such cases have grown exponentially, presumably for two reasons.[45] First, the databases are ever-expanding. In both the United States and Britain, state and federal DNA databases have now grown to include not only convicted felons, but also misdemeanants, juveniles, and even arrestees.[46] There are now over 7.8 mil-

---

[43] *See, e.g.*, Andrea L. Roth & Edward J. Ungvarsky, *Data Sharing in Forensic Science: Consequences for the Legal System*, 2009 AM. STAT. ASS'N, PROC. JOINT STAT. MEETING 469, 470–71 (discussing need for outside assessment of government claims); Erin Murphy, *Give Scholars Access to the National DNA Database*, S.F. CHRON., Feb. 24, 2009, at A11 (arguing that outside review is needed based on "broad reliance on the accuracy of DNA matches"); D.E. Krane et al., *Time for DNA Disclosure*, 326 SCI. 1631, 1631–32 (2009) (calling for greater scholarly access to DNA databases); *see also* Song et al., *supra* note 32, at 22, 24 (noting justice concerns about increased use of cold hits and calling for continued testing of assumptions behind government model). Even Professor David Kaye, who predicts that researchers will likely not find anything in the FBI's databases to contradict the FBI's RMP estimates, has called upon the FBI to allow access and end the controversy once and for all. David H. Kaye, *Trawling DNA Databases for Partial Matches: What Is the FBI Afraid of?*, 19 CORNELL J.L. & PUB. POL'Y 145, 170–71 (2009) (encouraging government to make anonymized version of data available to researchers to increase confidence in use of data).

[44] DNA Identification Act of 1994, Pub. L. No. 103-322, §§ 210,301–06, 108 Stat. 1796, 2065–71 (codified as amended in scattered sections of 42 U.S.C.), authorized the FBI to establish a national DNA database for identification purposes. The FBI's National DNA Index System (NDIS) became operational in October 1998, and the Combined DNA Index System (CODIS), enables information exchanges between NDIS, state, and local authorities. *See* FED. BUREAU OF INVESTIGATION, U.S. DEP'T OF JUSTICE, THE FBI'S COMBINED DNA INDEX SYSTEM PROGRAM—CODIS 1 (2000), *available at* http://www.rootsecure.net/content/downloads/pdf/fbi_codis.pdf.

[45] *Compare* NAT'L RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE 134 (1996) (noting that "more than 20 suspects have been identified" through offender database searches), *with* Fed. Bureau of Investigation, CODIS—NDIS Statistics, http://www.fbi.gov/hq/lab/codis/clickmap.htm (last visited Aug. 16, 2010) [hereinafter CODIS—NDIS Statistics] ("As of June 2010, CODIS has produced over 120,300 hits assisting in more than 117,800 investigations.").

[46] As of last year, thirteen states collected DNA samples from arrestees. In addition, federal arrestees are now subject to DNA testing regardless of whether they are ultimately convicted of the charged offense. Randal C. Archibold, *Justice Department Details Program for Collecting DNA from People in Federal Custody*, N.Y. TIMES, Apr. 19, 2008, at A11. And in May 2010, the House of Representatives passed a bill providing financial incentives to states that collect DNA from arrestees for certain serious crimes. Katie Sepich Enhanced DNA Collection Act of 2010, H.R. 4641, 111th Cong. (2010). The United Kingdom now permits DNA sampling of arrestees for recordable offenses. *See* CAROLE MCCARTNEY, FORENSIC IDENTIFICATION AND CRIMINAL JUSTICE 195 (2006) (discussing

lion offender profiles in the FBI's Combined DNA Index System (CODIS), a group of federal and state databases that together have produced over 120,300 cold hits.[47] Second, localities have received significant federal grant money to resolve backlogged cases using DNA evidence. In 2003, in an effort to expand and improve the use of DNA for solving crimes, then-President George W. Bush launched a $1 billion "DNA Initiative" to fund state agencies' examination of over 100,000 DNA cases.[48] As of 2009, more than 5000 database hits had resulted from the DNA Initiative alone.[49]

The number of "pure cold hit" DNA cases is also expanding. In those cases, the entirety of the government's case—other than the prior interaction with legal authorities that placed the suspect in the database—consists of the DNA match and perhaps general evidence, such as the perpetrator's race or age, that limits the potential suspect population only to a small and quantifiable degree. In some cases, the reviewing court has described the case before it as involving only DNA match evidence.[50] In others, the DNA match is accompanied

---

change in U.K. practices). Also, as of 2008, twenty-eight states collected DNA samples from juveniles adjudicated delinquent. Natalie A. Bennett, *A Privacy Review of DNA Databases*, 4 I/S: J.L. & POL'Y FOR INFO. SOC'Y 821, 840 (2008). Indeed, some politicians and legal scholars have called for the establishment of population-wide citizen DNA databases. *See id.* at 837–38.

[47] *See* CODIS—NDIS Statistics, *supra* note 45 (describing success of CODIS program).

[48] *See* NAT'L RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 71–72 (2009) (describing initiative).

[49] *Id.* at 72.

[50] *See, e.g.*, State v. Hunter, 861 N.E.2d 898, 901 (Ohio Ct. App. 2006) (noting in rape case that "*literally no other evidence linked appellant to the crime*" other than DNA match with RMP of 1 in 756 trillion (emphasis added)); State v. Toomes, 191 S.W.3d 122, 129 (Tenn. Crim. App. 2005) ("The only—and we emphasize *the only*—evidence connecting [Toomes] to the victim's rape is the DNA results."); R v. Lashley [2000] EWCA (Crim) 88, [3] (Eng.) (describing jury instruction explaining that Lashley was 1 of 7–10 males in United Kingdom expected to share this DNA profile, and noting that jury convicted even though there was no other evidence against defendant); R v. Adams, [1996] 2 Crim. App. 467, 469–70 (Eng.) (finding no inculpatory evidence against suspect except DNA match with disputed RMP between 1 in 2 million and 1 in 200 million); *see also* Mike Redmayne, *Rationality, Naturalism, and Evidence Law*, 2003 MICH. ST. L. REV. 849, 879–80 [hereinafter Redmayne, *Rationality*] (discussing *Lashley* and noting that RMP was 1 in 4 million); Mike Redmayne, *Appeals to Reason*, 65 MOD. L. REV. 19, 21 (2002) (discussing *Lashley*); *cf.* State v. Davis, 698 N.W.2d 823, 826–27 (Wis. Ct. App. 2005) (describing evidence for warrant issued against Davis, accused of 1994 rape, as being cold hit found in 1998–1999 and laboratory's conclusion that "only reasonable scientific explanation" for match "was that Davis was the source of the semen in [the victim's] underwear"). Other apparently "pure cold hit" cases have not yet been reviewed by an appellate court. *See, e.g.*, Preliminary Hearing Minutes, People v. Moore, No. YA062921 (Cal. Super. Ct. May 30, 2007) (basing case against defendant solely on DNA match with RMP with denominator in quintillions). Of course, even in those cases described by courts as having only DNA match evidence, undisputed facts presumably exist that would narrow the potential suspect popu-

only by a general description of the perpetrator's physical appearance or his connection to a large geographic area that only moderately limits the size of the suspect population.[51] The RMPs in existing pure cold hit cases range from 1 in 1.1 million[52] to those with denominators in the quintillions or beyond.[53]

---

lation. *See* Kaye, *supra* note 6, at 487 (noting in hypothetical "naked trawl" case that "some men would have been in the hospital or have been too feeble to have carried [the victim] away; many would have airtight alibis; many would be psychologically incapable of the alleged conduct").

[51] *See, e.g.*, State v. Raines, 857 A.2d 19, 21–22 (Md. 2004) (stating that only evidence was that both perpetrator and Raines were black males of medium height); Roberson v. State, 16 S.W.3d 156, 160–63 (Tex. App. 2000) (providing no evidence except cold hit and match to blood type common to 9–10% of all men); R v. Watters [2000] EWCA (Crim) 89 (noting no inculpatory evidence except DNA match linking defendant to cigarette butts found at five apparently related burglaries in Birmingham, England, evidence that defendant was smoker and that he lived in Birmingham, and fact that he was male, coupled with assumption that most people who crack safes are men); Appellant's Opening Brief, *supra* note 1, at 38, 56–57, 102–03 (describing primary evidence as cold hit with RMP of 1 in 1.1 million, prior sexual assault conviction, and Puckett's connection to Bay Area); Official Transcript of Proceedings at 4-12, State v. Derr, No. 08-K-04-000930 (Md. Cir. Ct. Aug. 31, 2006) (recording defendant's motion for acquittal based on evidentiary insufficiency since only evidence identifying defendant was cold hit and very general physical description of perpetrator). In other cases, the additional information is still general but limits the suspect population more narrowly to a neighborhood or those seen around the crime area. *See, e.g.*, Riggs v. State, 809 N.E.2d 322, 324 (Ind. 2004) (mentioning that only evidence against Riggs for 1985 rape-murder was 2000 cold hit and fact that Riggs lived near location where victim went missing); State v. Abdelmalik, 273 S.W.3d 61, 63 (Mo. Ct. App. 2008) (describing no inculpatory evidence except DNA match with RMP of 1 in 1 quintillion and fact that defendant lived within two miles and worked within one mile of crime scene at time of 1980 killing); People v. Rush, 630 N.Y.S.2d 631, 632 (Sup. Ct. 1995), *aff'd*, 672 N.Y.S.2d 362 (App. Div. 1998) (noting that only evidence permissible in court was cold hit and fact that defendant had been seen in "area" of robbery-rape three days before incident); *cf.* People v. Soto, 981 P.2d 958, 961 (Cal. 1999) (describing only inculpatory evidence as DNA match with RMP of 1 in 189 million in Hispanic population and fact that victim thought masked rapist's voice might be defendant's because defendant had come over earlier that day, while noting that defendant (Latino, black hair, dark complexion) did not match victim's physical description of rapist (white, blond hair, olive complexion)); Springfield v. State, 860 P.2d 435, 449 (Wyo. 1993) (providing no evidence other than DNA match, fact that defendant was consistent with general description of perpetrator, and that victim said at trial that defendant "definitely resembles" her attacker, though she could not identify him).

[52] *See, e.g.*, Appellant's Opening Brief, *supra* note 1, at 38 (describing prosecution's case in *Puckett*, including its assessment of defendant's RMP).

[53] *See, e.g.*, *Abdelmalik*, 273 S.W.3d at 66 (relying on RMP of 1 in 1 quintillion). None of these cases has discussed what the source probability might be, given the RMP and the population of likely suspects. Instead, nearly all of these courts have engaged in the fallacy of the transposed conditional, *see infra* note 95 and accompanying text (distinguishing between RMP and source probability), and treated the RMP as the source probability. While low RMPs on the order of 1 in 1 quintillion necessarily imply astronomically high source probabilities as well, higher RMPs may well be consistent with a source probability of less than 99.9%. *See infra* Part V (examining cases discussed for whether they would be resolved at sufficiency stage because of high RMP).

Although some pure cold hit cases have ended in post-trial conviction and affirmance on appeal after trial,[54] others have ended in pleas,[55] acquittals,[56] or reversal on appeal.[57] In addition, several pure cold hit suspects have been exonerated upon discovery by law enforcement that the cold hit was coincidental[58] or the likely product of laboratory contamination[59] or interpretive errors.[60]

---

[54] *See, e.g.*, *Hunter*, 861 N.E.2d at 901 (affirming conviction after trial though "*literally no other evidence linked appellant to the crime*" other than DNA match with RMP of 1 in 756 trillion (emphasis added)); *Toomes*, 191 S.W.3d at 129 (affirming conviction after trial though "[t]he only—and we emphasize *the only*—evidence connecting [Toomes] to the victim's rape is the DNA results").

[55] *See, e.g.*, Motion To Compel Government To Produce Statistics on Matching and Near-Matching Profiles Maintained in the CODIS DNA Database at 1, United States v. Rue, Crim. No. F-3817-05 (D.C. Super. Ct. Dec. 3, 2005) ("The sole basis for the indictment of Mr. Rue was a DNA database 'cold hit.' . . . Prior and subsequent to the DNA search, to the best knowledge of defense counsel, no other evidence has been alleged to link Mr. Rue to the complaining witness."); Response to Defendant's Motion To Compel Production of Statistics on Matching and Near-Matching Profiles Maintained in the CODIS Database at 2, United States v. Rue, Crim. No. F-3817-05 (D.C. Super. Ct. Jan. 6, 2006) (referring to match between defendant and DNA sample as "cold hit"). Rue entered a plea of guilty on January 11, 2006. *See* Court File, United States v. Rue, Crim. No. F-3817-05 (D.C. Super. Ct. Jan. 11, 2006).

[56] Tim O'Neil & William C. Lhotka, *DNA Is Just a Start in Two Old Murders: Expert Says He Doesn't Know of a Single Rape or Murder Conviction Based on DNA Alone*, Sт. Louis Post-Dispatch, Feb. 13, 2007, at A1 (noting acquittal of rape defendant on twenty-seven counts resting on DNA evidence alone).

[57] *See, e.g.*, R v. Lashley, [2000] EWCA 88, [3], [14]–[16] (setting aside conviction based on insufficiency of evidence when DNA could have matched "five or six men in the United Kingdom"); *see also* Redmayne, *Rationality*, *supra* note 50, at 880 (discussing *Lashley*).

[58] For example, police investigating a burglary in Bolton, England, recovered blood off a window and got a cold hit match to Raymond Easton, a man who lived two hundred miles away and was in the British database by virtue of a previous domestic dispute. Though Easton could barely walk because of Parkinson's disease and had a corroborated alibi, police insisted he was the culprit based on the RMP of 1 in 37 million. Easton was exonerated after police conducted more advanced DNA testing at additional loci. Cole & Lynch, *supra* note 7, at 48; *see also* William C. Thompson, *The Potential for Error in Forensic DNA Testing (and How That Complicates the Use of DNA Databases for Criminal Identification)*, 2008 Council for Responsible Genetics Nat'l Conf. 9, http://www.councilforresponsiblegenetics.org/pageDocuments/H4T5EOYUZI.pdf (noting coincidental cold hit match between two unrelated crime scene DNA samples); *id.* at 33 (noting cold hit tying Nottingham boy to Northern Ireland terrorist bombing case that appeared to be coincidental match).

[59] For example, police investigating a 1969 Michigan murder recently conducted DNA testing on stains on the victim's clothing, which yielded database hits to two men: Gary Leiterman, a sixty-two-year-old nurse with a forgery conviction, and John Ruelas, a forty-year-old convicted murderer. People v. Leiterman, No. 265821, slip op. at 2 (Mich. Ct. App. July 24, 2007). Though Ruelas would have been only four years old at the time of the murder, prosecutors insisted he was a "chronic nose-bleeder" and, while not the murderer, had bled at the scene. Cole & Lynch, *supra* note 7, at 48. A more likely explanation for the hit to Ruelas is contamination; both men's samples were processed in the same laboratory at the same time as the sample from the 1969 case. *Id.*; *see also* William C. Thompson, *Tarnish on the "Gold Standard": Understanding Recent Problems in Forensic DNA Testing*,

Several factors suggest that pure cold hit cases may grow significantly in the coming years. First, the federal government is giving states grant money to solve closed "no suspect" cases, some from twenty or thirty years ago.[61] In these old cases, witnesses may be unable or unwilling to give testimony[62]—indeed, numerous key witnesses had died by the time of Mr. Puckett's trial[63]—and the government may lack a coherent theory of motive or any non-DNA forensic evidence.[64] The modern practice in some unsolved cases of indicting a DNA profile with the name "John Doe" when police have an evidence sample but not a suspect, to save the prosecution from statutes of limitations, will continue to make such old cases possible.[65] Second, the very discovery of a cold hit may cause police to feel that additional time-consuming and expensive factual investigation is unnecessary.[66]

CHAMPION, Jan./Feb. 2006, at 10, 13 (discussing Washington "false 'cold hit[ ]'" where juvenile respondent's database sample was used as "training sample" in laboratory at same time another kit was analyzed for rape committed when juvenile was a young child); *id.* at 13–14 (discussing Australian toddler-murder case where young mentally disabled woman, herself a victim of sexual assault, was falsely inculpated by cold hit when her own DNA, found on outside of condom used in sexual assault against her, was analyzed at same time as toddler's clothing); Thompson, *supra* note 58, at 9 (noting coincidental cold hit in Chicago burglary case to woman incarcerated at time of offense); *id.* at 25 (citing examples of false cold hits due to likely contamination between samples); *id.* at 28–29 (discussing admitted false cold hit due to suspected contamination in New Jersey case).

[60] *See, e.g.*, Thompson, *supra* note 58, at 31 (noting Sacramento rape case in which police got cold hit match to man who lived in local area and when further investigation suggested his innocence, laboratory reviewed analyst's work and found interpretive error); *cf. id.* at 33 (discussing near-exclusion in 2005 of Missouri man whose CODIS profile matched evidence sample at twelve of thirteen loci, when police realized that man's profile had been mislabeled at non-matching locus and in fact all thirteen loci matched).

[61] *See supra* notes 48–49 and accompanying text (discussing DNA Initiative).

[62] *See, e.g.*, Song et al., *supra* note 32, at 22 ("[C]old-hit cases raise justice-related concerns, especially since mounting a defense to a crime that occurred in the past becomes increasingly difficult as time progresses.").

[63] Appellant's Opening Brief, *supra* note 1, at 4 & n.6, 9 (noting that victim's landlady, investigating police officers, and medical examiners were all dead or incapacitated by time of trial).

[64] *Cf.* Solomon Moore, *Progress Is Minimal in Clearing DNA Cases*, N.Y. TIMES, Oct. 25, 2008, at A9 ("An audit . . . found that 217 backlogged cases [in Los Angeles] involved sexual assaults so old the 10-year statute of limitations had lapsed."); Cole & Lynch, *supra* note 7, at 46 (noting high number of backlogged and older cases waiting for analysis and database searches).

[65] *See generally* Meredith A. Bieber, *Meeting the Statute or Beating It: Using "John Doe" Indictments Based on DNA To Meet the Statute of Limitations*, 150 U. PA. L. REV. 1079 (2002) (describing new practice of issuing "John Doe" indictments to allow police to "continue their efforts to identify the suspect").

[66] *See, e.g.*, McCARTNEY, *supra* note 46, at 63, 183 (discussing concern that "faith in forensic science has been too easily used to shore up falling confidence in police investigative competence" and explaining that once investigators recover fingerprints or DNA from crime scene, both police and defense attorneys tend to neglect follow-up investigation); Cole & Lynch, *supra* note 7, at 44 ("Forensic investigation is positioned to displace much of

Third, there appears to be a trend in law enforcement toward reliance on DNA rather than other lesser forms of direct but unreliable evidence such as eyewitness identification.[67] The highly publicized use of DNA typing as a tool of exoneration in cases involving wrongful convictions based on such lesser forms of evidence[68] has likely fueled the perception that DNA evidence, when present, is dispositive.[69]

The increasing frequency of pure cold hit cases—including false cold hits ending in exoneration—has not escaped the notice of journalists and scholars.[70] Some have voiced concern that DNA alone carries too great a risk of laboratory error,[71] coincidental matches,[72] or

the 'good, old-fashioned detective work' of canvasing [sic] neighborhoods, questioning witnesses, interrogating suspects, and cultivating informants." (quoting MARK FUHRMAN, DEATH AND JUSTICE 222 (2003))).

[67] *See* MCCARTNEY, *supra* note 46, at 32 (quoting official as acknowledging that: "You are not going to get admissions [from suspects] these days, the use of informants is not as effective as it was, even societal factors such as the public not relating to the police as they once did . . . as other things have fallen away, forensic evidence is what is left").

[68] *See, e.g.*, Adam Liptak, *Study of Wrongful Convictions Raises Questions Beyond DNA*, N.Y. TIMES, July 23, 2007, at A1 (noting that two hundred prisoners have been shown innocent because of DNA evidence); The Innocence Project, http://www.innocenceproject.org/know/Browse-Profiles.php (last visited July 26, 2010) (listing and linking to information on over two hundred cases in which convicted defendants were exonerated through DNA testing).

[69] Given the clear differences in using DNA as a tool of exclusion versus inclusion, this perception may not be justified. *See* discussion *infra* note 145 (discussing differences between DNA exoneration and DNA proof of guilt).

[70] *See, e.g.*, MICHAEL LYNCH ET AL., TRUTH MACHINE: THE CONTENTIOUS HISTORY OF DNA FINGERPRINTING 184 (2008) ("[I]n some cases the cold hit is the sole, or main, item of criminal evidence."); Cole & Lynch, *supra* note 7, at 47 (observing "some very interesting cases" in which DNA "was not corroborated and was even contradicted by other evidence"); Kaye, *supra* note 6, at 490–92 (discussing *Puckett* as database trawl case based on defendant's regional proximity); Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CAL. L. REV. 721, 741 (2007) ("[I]n some cases, the government may proceed on the sole basis of genetic evidence or marginally probative additional evidence, such as the suspect's proximity to the scene of the offense."); Song et al., *supra* note 32, at 22 ("Cold-hit cases have prompted courts to confront the question of whether a genetic match constitutes sufficient evidence to uphold a conviction.").

[71] *See generally* Boaz Sangero & Mordechai Halpert, *Why a Conviction Should Not Be Based on a Single Piece of Evidence: A Proposal for Reform*, 48 JURIMETRICS J. 43, 45 (2007) ("[T]he possibility of a laboratory error . . . is much more likely than a random match."); Brooke G. Malcom, Comment, *Convictions Predicated on DNA Evidence Alone: How Reliable Evidence Became Infallible*, 38 CUMB. L. REV. 313 (2008) (discussing concerns about human error and fraud with regard to forensic testing in laboratories).

[72] *See, e.g.*, Song et al., *supra* note 32, at 22–24 (discussing probability of erroneous cold hit matches); Jason Felch & Maura Dolan, *DNA: Genes as Evidence*, L.A. TIMES, July 20, 2008, at A1 (noting problems with using matches obtained through database searches due to potential for incorrect matches); Humes, *supra* note 3, at 20–22 (using *Puckett* as example where possibility exists of coincidental match); Smith, *supra* note 2 (describing existence of data suggesting that coincidental "low-level matches" are much more common

other injustice[73] to be permissible. Nevertheless, it seems clear that pure DNA-based prosecutions, because of their proffered near-certain statistical proof of guilt, are "well on the way to remaking the technology of justice."[74]

II

THE LEGAL SUFFICIENCY REQUIREMENT AND ITS PROBLEMATIC APPLICATION IN PURE COLD HIT CASES THUS FAR

A criminal defendant has a due process right not to be tried on legally insufficient evidence.[75] Evidence against a defendant is insufficient as a matter of law if no rational juror could find that the evidence proves the defendant's guilt beyond a reasonable doubt.[76] The trial judge typically makes this determination upon the defendant's motions for judgment of acquittal both at the close of the government's case and at the close of all the evidence.[77] Viewing the evidence in the light most favorable to the government, the non-moving party, the judge must grant the defendant's motion if the judge determines that no rational juror could find the evidence to constitute proof beyond a reasonable doubt.[78] The purpose of the sufficiency

---

than previously thought); Thompson, *supra* note 58, at 10 (suggesting probability of coincidental matches is higher from cold hits than from profile matching).

[73] *See, e.g.*, MCCARTNEY, *supra* note 46, at 101 (arguing that courts "need to demand independent corroborative evidence" before convicting defendant on basis of forensic identity evidence alone, though not specifying why); *id.* at 145 (reporting that Sir Alec Jeffries, the father of modern DNA testing, had qualified his early support for comprehensive national DNA database in Britain by stating that cold hit alone should not be considered conclusive evidence of guilt and had expressed concern about potential discrimination in database samples); NUFFIELD REPORT, *supra* note 26, at 71 (arguing that because misleading DNA evidence could be given great weight in courtroom, it is "vital that defendants are not convicted on a DNA match alone").

[74] Cole & Lynch, *supra* note 7, at 44; *see also* Murphy, *supra* note 70, at 723 (describing "new generation of forensic sciences," including DNA typing, as "stak[ing] a central and indispensable role in the future administration of criminal justice").

[75] *See* Jackson v. Virginia, 443 U.S. 307, 316 (1979) (describing reasonable doubt standard as necessary to protect due process of law); *see also* Herrera v. Collins, 506 U.S. 390, 402 (1993) (reaffirming that "the *Jackson* inquiry is aimed at determining whether there has been an independent constitutional violation—i.e., a conviction based on evidence that fails to meet the *Winship* standard," requiring proof beyond reasonable doubt on each element of offense as matter of due process).

[76] *See Jackson*, 443 U.S. at 320 (holding that sufficiency requires more than "mere modicum" of evidence).

[77] *See, e.g.*, Richardson v. United States, 468 U.S. 317, 318 (1984) (citing instances of defendant's motions for acquittal).

[78] *See, e.g.*, *Jackson*, 443 U.S. at 324 (articulating standard for habeas relief based on insufficiency of evidence); *see also* WAYNE R. LAFAVE ET AL., CRIMINAL PROCEDURE § 24.6(c) (3d ed. 2000) (describing sufficiency standard).

requirement is to ensure that a defendant is never convicted on less than proof "beyond a reasonable doubt."[79]

When a criminal case involves qualitative evidence such as an eyewitness identification or confession, a judge cannot easily determine whether no rational juror would find the evidence to be proof beyond a reasonable doubt. The assessment of such evidence turns on each juror's determination of the credibility of witnesses, and short-circuiting the prosecution before trial would presumably usurp the credibility-finding role of the jury. Courts typically dispose of qualitative cases on sufficiency grounds only when the evidence fails to meet a clear statutory or common law requirement, such as failure to prove the operability of a weapon,[80] failure to comply with the corpus delicti rule prohibiting conviction based solely on an uncorroborated confession,[81] or the fact that the instrument involved in a crime does not meet the definition of a "deadly weapon."[82] Thus, in cases involving both DNA evidence and qualitative corroborative evidence of guilt, judges typically allow the case to go to the jury, and sufficiency issues related to DNA evidence are rare (if they exist at all) on appeal.[83]

In a case where the only evidence is a DNA profile match, however, a trial judge could conceivably determine mathematically,

---

[79] *Jackson*, 443 U.S. at 320 n.14; *see also* Keith A. Findley & Michael S. Scott, *The Multiple Dimensions of Tunnel Vision in Criminal Cases*, 2006 Wis. L. Rev. 291, 348 (noting that *Jackson* ruling protected "constitutional requirement of proof beyond a reasonable doubt").

[80] *See, e.g.*, People v. Brun, 872 N.Y.S.2d 188, 191 (App. Div. 2009) (reversing conviction for "criminal use of a firearm" on sufficiency grounds where government failed to prove firearm was operable, which is one element of offense).

[81] *See, e.g.*, Goodsell v. State, 289 S.W.3d 534, 535–36 (Ark. Ct. App. 2008) (reversing sexual assault conviction for insufficient evidence where alleged victim recanted prior testimony and government failed to corroborate defendant's confession with other substantive evidence of assault).

[82] *See, e.g.*, Berry v. State, 212 P.3d 1085, 1094 (Nev. 2009) (holding that government failed to prove that pellet gun met definition of "deadly weapon").

[83] In nearly all cases I have found in which an appellant raised a sufficiency claim related to DNA evidence, the appellant's argument was that the DNA is uncorroborated and insufficient on its own. For examples, see *infra* note 90. In one case, a litigant argued that, notwithstanding corroborative evidence of guilt, DNA evidence was prejudicial because it turned the case into the type of "trial by mathematics" condemned in *People v. Collins*, 438 P.2d 33, 38–42 (Cal. 1968). *See* United States v. Wright, 215 F.3d 1020, 1028 (9th Cir. 2000) (rejecting appellant's argument on grounds that "*Collins* is not a DNA case" and is "simply inapplicable"). In *Collins*, the court reversed the conviction of one of the defendants due to prejudicial testimony given by a mathematician who improperly used the product rule to estimate the probability that the defendants were guilty based on various characteristics they shared with the victim's and a witness's description of the suspects. 438 P.2d 33. To bring an argument in a pure cold hit case based on *Collins*, a defendant would have to argue that the match statistic as calculated was somehow unreliable because of a lack of independence among variables. Should future studies show that currently reported RMPs are indeed unreliable, such arguments may gain traction.

without usurping the jury's credibility-determining role, the extent to which the evidence tends to show a high or low "source probability," that is, the probability that the defendant, and not some other person, is the source of the DNA. Where a DNA profile match is the only evidence against a suspect, a court can calculate the source probability from the RMP and the size of the suspect population by using Bayes' Theorem, a common and well-established mathematical formula.[84] For example, if the suspect population were the number of men in California (approximately 18 million),[85] the initial or "prior" odds of guilt would be about 18 million to 1.[86] If the RMP were 1 in 180 million (meaning that we would expect 1 in every 180 million people to have the profile), and the government offers no other admissible inculpatory evidence of guilt, the best that the government would be able to offer the jury is about 10-to-1 odds, or about a 91% chance, that the defendant is the source of the DNA based on the profile match.[87]

---

[84] Scholars appear to be in agreement on this point. *See, e.g.*, David J. Balding & Peter Donnelly, *How Convincing Is DNA Evidence?*, 368 NATURE 285, 285–86 (1994) (using version of Bayes' Theorem to illustrate calculation of source probability); Kaye, *supra* note 6, at 491–92 (explaining how to determine source probability using simplified version of Bayes' Theorem); Thompson, *supra* note 58, at 11 n.2 (noting that size of unrelated population multiplied by probability of finding at least one additional individual with same genetic profile gives simplified estimate of chance of finding at least one match). Professor Richard Friedman offers a straightforward statement for the non-mathematician regarding the relevant terms used in Bayesian analysis:

> A simple statement of Bayes' Theorem uses three terms. One is the *prior odds* of a proposition—that is, the odds as assessed before receipt of the new evidence. The second is the *posterior odds* of the proposition—that is, the odds that the proposition is true as assessed after receipt of the new evidence. And the third is the *likelihood ratio*. Simply defined, the likelihood ratio of a given body of evidence with respect to a given proposition is the ratio of the probability that the evidence would arise given that the proposition is true to the probability that the evidence would arise given that the proposition is false.

Richard D. Friedman, Commentary, *A Presumption of Innocence, Not of Even Odds*, 52 STAN. L. REV. 873, 875 (2000).

[85] California Quick Facts from the U.S. Census Bureau, http://quickfacts.census.gov/qfd/states/06000.html (last visited July 24, 2010).

[86] Absent non-DNA evidence of guilt, our initial estimate of the chance that the defendant is the source is the inverse of the suspect population.

[87] Using Bayes' Theorem, the odds that the defendant is the source (or posterior odds) would be the prior odds that the defendant is the source multiplied by the likelihood ratio, which is the inverse of the RMP—that is, the chance of the match assuming the defendant is the source (one), divided by the chance of the defendant matching, assuming he is not the source (the RMP). In the example above, the posterior odds would be:

Prior Odds		Likelihood Ratio		Posterior Odds
$\dfrac{1}{18\text{ million}}$	$\times$	$\dfrac{180\text{ million}}{1}$	$=$	$\dfrac{10}{1}$

A court would then presumably have to decide whether a 91% chance that the defendant is the source of the DNA is high enough to justify the legal conclusion that the DNA match statistic alone is legally sufficient evidence of guilt. Of course, the source probability is different from the ultimate probability that the defendant is guilty of the charged offense.[88] Nonquantifiable factors might still influence the jury's ultimate determination of whether the defendant is guilty, including legal defenses, such as innocent presence or consent, that do not contradict the DNA evidence.[89] Any other exculpatory evidence, such as a persuasive alibi or the victim's description of her attacker as a tall Caucasian man (when the defendant is a short African-American man), might also sway a jury. But these factors would, if anything, decrease the chance of guilt to a point lower than the 91% source probability. Faced with an uncorroborated DNA match statistic that, at most, shows a 91% upper bound on the chance of the defendant's guilt, the court would be forced to decide whether that number is high enough to constitute proof beyond a reasonable doubt.

A number of litigants have attempted to argue in pure cold hit cases that a DNA profile match alone, uncorroborated by other individualized evidence, is legally insufficient proof of guilt.[90] Most of these litigants have argued that a DNA match, like blood typing, at worst places the defendant in a pool of potential suspects rather than

---

While the posterior "odds" express the ratio of the probability the defendant is the source to the probability he is not, the posterior "probability" would be equal to (odds)/(odds+1) = 10/11 = 91%. *See* NAT'L RESEARCH COUNCIL, *supra* note 45, at 132 (1996) (explaining conversion of odds to probability when using Bayes' Theorem). Note that this calculation assumes a uniform prior probability throughout the suspect population. While this is not entirely realistic, since various persons in the suspect population will be more or less likely than the defendant to be the source of the DNA, the government will likely be unable to prove more favorable prior odds without information about those in the suspect population. *See* Kaye, *supra* note 6, at 492 (noting assumptions underlying this "simplified version" of Bayes' Theorem); Song et al., *supra* note 32, at 24 (noting that while assumption that "*a priori*, each individual in the population is equally likely to have committed the crime" is not precise, it is acceptable estimate to make results "conservative"); Thompson, *supra* note 58, at 11 n.2 (noting possibility of using uniform prior probability distribution among suspect population). Note also that this formula assumes a zero false-positive rate in the DNA-typing process.

[88] *See, e.g.*, McDaniel v. Brown, 130 S. Ct. 665, 670 (2010) ("It is further error to equate source probability with probability of guilt, unless there is no explanation other than guilt for a person to be the source of crime-scene DNA.").

[89] As noted previously, the nature of the evidence in most pure cold hit cases makes these defenses untenable. *See supra* note 32 (discussing sexual assault cases in which type of evidence available forecloses possibility of innocent presence).

[90] *See* People v. Soto, 35 Cal. Rptr. 2d 846, 859 (Ct. App. 1994), *aff'd*, 981 P.2d 958 (Cal. 1999) (describing defendants' attempts to argue that cold hit is legally insufficient); State v. Abdelmalik, 273 S.W.3d 61, 65 (Mo. Ct. App. 2008) (same); People v. Rush, 630 N.Y.S.2d 631, 632 (Sup. Ct. 1995) (same); State v. Hunter, 861 N.E.2d 898, 901 (Ohio Ct. App. 2006) (same); Roberson v. State, 16 S.W.3d 156, 159 (Tex. App. 2000) (same).

conclusively identifying the defendant as the source of the evidence, like fingerprints ostensibly do.[91] Thus, these litigants seem to have implicitly acknowledged that the "source probability" is the ultimate question in determining sufficiency in a pure cold hit case. No litigant has yet argued that probabilistic evidence alone is inherently insufficient as a matter of law to prove guilt.[92]

Appellate courts in the United States have uniformly rejected these sufficiency challenges.[93] In doing so, as catalogued below, nearly all have relied on impressive-sounding RMPs—with denominators ranging from the millions to the quintillions—to conclusorily declare that the DNA evidence is sufficient because it shows beyond peradventure that the defendant is the source. Thus, these courts also appear to implicitly recognize that the source probability is the ultimate question in determining sufficiency in a pure cold hit case. Yet none has discussed how high a source probability would have to be for an uncorroborated DNA profile match to be legally sufficient, nor explained why an RMP with a denominator in the millions is legally sufficient proof that the defendant is the source of the DNA when the relevant suspect population may well include millions of people, making another matching suspect likely. If any court recognized that the chance of another suspect matching was a real possibility, it would presumably have to decide whether that possibility is great enough to create reasonable doubt as a matter of law.

One clear reason for American courts' failure to fully explore the exact point at which a match statistic becomes legally sufficient evidence of guilt is that nearly every one has succumbed to the "fallacy of

---

[91] *See Soto*, 35 Cal. Rptr. 2d at 859 (describing defendant's argument that DNA evidence should have to be corroborated because it only brought Soto "within a *class* of potential suspects"); *Abdelmalik*, 273 S.W.3d at 66 (describing defendant who argued that DNA match alone did not provide "substantial evidence of identity"); *Rush*, 630 N.Y.S.2d at 632 (recounting testimony of DNA expert who distinguished fingerprinting from DNA matching in terms of former's ability to result in "absolute identification," contrasted with latter's lack of ability to do so); *Roberson*, 16 S.W.3d at 159 (describing appellant who argued "that DNA evidence is not an inclusionary tool but one of probability," given that "DNA testing can exclude someone as the donor of semen, but it cannot definitively link someone to a crime without other evidence"). In *Hunter*, the appellant argued "that the jury clearly lost its way in giving credence to the DNA results when literally no other evidence linked appellant" to the charged stranger-rape. 861 N.E.2d at 901. His claim appeared to be not that the DNA was insufficient because the probability of a coincidental match was too great, but that the DNA was not sufficiently reliable because of alleged gaps in the chain of custody. *Id.*

[92] *Cf. infra* Part III.C (arguing that very high source probabilities can justify conviction).

[93] For cases rejecting sufficiency challenge, see, for example, *Abdelmalik*, 273 S.W.3d at 66, *Rush*, 630 N.Y.S.2d at 634, *Hunter*, 861 N.E.2d at 901, *State v. Toomes*, 191 S.W.3d 122, 131 (Tenn. Crim. App. 2005), and *Roberson*, 16 S.W.3d at 171.

the transposed conditional,"[94] conflating the RMP with the source probability. Recall that while the former is the chance that a random person from the population would match the profile (and is thus an estimate of the frequency of the profile in the population), the latter is the chance that the defendant is the source of the DNA given the evidence.[95] An RMP of 1 in 1000 does not signify that there is only a 1 in 1000 chance that someone other than the defendant is the source of the DNA. Rather, it means that a person randomly selected from the population has a 1 in 1000 chance of matching the profile, or, equivalently, that we would expect 1 in every 1000 people to share the profile.[96] In a population of 20,000 people, for example, we would expect about twenty people to match. Thus, the match alone only puts the defendant within a group of twenty possible sources, a far cry from suggesting only a 1 in 1000 chance that he might not be the source.

The fallacy takes center stage in nearly every discussion about sufficiency of evidence in a pure cold hit case in U.S. appellate courts. For example, the Tennessee Court of Appeals in *State v. Toomes*, a case in which "*the only* . . . evidence connecting [Toomes] to the

---

[94] *See* McDaniel v. Brown, 130 S. Ct. 665, 670 (2010) (acknowledging this fallacy as "the prosecutor's fallacy"); Brief of 20 Scholars of Forensic Evidence as *Amici Curiae* in Support of Respondent, at 12 & nn.7–9, McDaniel v. Brown, 130 S. Ct. 665 (2010) (No. 08-559) [hereinafter *McDaniel* Amici Brief] (noting frequency of "fallacy of the transposed conditional" and difference between that fallacy and similar "prosecutor's fallacy," where government mistakes RMP for chance that defendant is innocent); William C. Thompson, Letter to the Editor, *The Prosecutor's Fallacy in George Clarke's* Justice and Science: Trials and Triumphs of DNA Evidence, 54 J. FORENSIC SCI. 504, 504 (2009) (noting literature documenting "well-known prosecutor's fallacy" and pointing out Judge Clarke's own fallacious statements conflating RMP and chance that observed match is coincidental).

[95] The tendency of courts, litigants, and journalists to misunderstand the RMP and source probability may also be related to their near-exclusive focus in the last few years on the so-called "database search controversy," a dispute among statisticians as to how best to communicate the strength of cold hit evidence to a jury by using the RMP and/or the chance of finding at least one match in the offender database searched (the "database match probability," or DMP). *See, e.g.*, People v. Nelson, 185 P.3d 49, 62–64 (Cal. 2008) (recounting four different methods of analyzing statistical significance of match); United States v. Jenkins, 887 A.2d 1013, 1019–20 (D.C. 2005) (describing "raging debate" among scientists over "which method of probability determination is the most significant for expressing a cold hit"). Ultimately, it is the chance of finding the profile in the potential suspect population, not the offender database, that is the relevant probability for determining the probative value of a cold hit. The concepts of the RMP and DMP (approximately the RMP multiplied by the size of the database), and the difference between those two probabilities, are not well understood by litigants. *See generally* David H. Kaye, Case Comment, People v. Nelson: *A Tale of Two Statistics*, 7 LAW PROBABILITY & RISK 249 (2008) (portraying *Nelson* court's difficulty in comprehending difference in probative value of DMP versus RMP). Ironically, in a pure cold hit case, the DMP may well be a less favorable statistic to the defendant than the source probability.

[96] *See McDaniel* Amici Brief, *supra* note 94, at 14–15 (distingushing between RMP and source probability by noting that "no meaningful conclusions" about source probability may be drawn "from DNA evidence alone").

victim's rape [wa]s the DNA results,"[97] concluded that the DNA profile match "was sufficient" evidence of guilt because the "probabilit[y] of finding the defendant's profile within the African-American population was 1 in 5 billion, 128 million," and was even smaller for other ethnic groups.[98] The court reached this conclusion even though the defense never argued this point on appeal.[99] But this conclusion is fallacious. If the relevant African-American suspect population were 5 million, then the chance of matching a suspect in that population to a profile with an RMP of 1 in 5 billion would be approximately 1 in 1000, not 1 in 5 billion. While the court might have still viewed a 1-in-1000 chance of another suspect as being small enough to justify conviction, the court—as a result of its mathematical error—did not grapple with that question.

In *State v. Hunter*, the Ohio Court of Appeals reasoned in a stranger-rape case that it "cannot conclude that the jury lost its way" where the "probability of the [semen] sample . . . belonging to anyone other than appellant was 1 in 756 trillion."[100] Again, 1 in 756 trillion is not the probability that the DNA belonged to someone other than the appellant; it is the probability that any randomly selected person from the population would match the profile. Of course, such a small RMP would yield an impressive source probability regardless of the size of the suspect population, given that the Earth's population is only six billion people. But the court's reasoning would presumably be equally flawed in a subsequent case with a higher RMP.

Other courts have committed similar errors in their sufficiency analyses. In *Roberson v. State*, a stranger-rape/burglary case, a Texas court held without analysis that a blood profile match, found in nine to ten percent of the male population, along with a DNA cold hit match, was sufficient evidence of guilt.[101] In doing so, the court inaccurately stated that the "chance of appellant's DNA profiling occurring in another person" was the RMP, 1 in 420 million for the African-American population.[102] In *People v. Soto*, the court rejected

---

[97] 191 S.W.3d 122, 129 (Tenn. Crim. App. 2005).

[98] *Id.* at 131. The court noted that it was not "announcing an iron-clad principal [sic] that DNA evidence, without corroboration, is always sufficient to support a conviction," and noted the "unusual feature" in Toomes's case that "two separate DNA comparisons were performed," one of Toomes and one of his twin brother, the man first implicated by the database cold hit. *Id.* at 131 n.4.

[99] *See* Brief of Appellant Darrell Toomes at 4–6, State v. Toomes, 191 S.W.3d 122 (Tenn. Ct. App. 2005) (No. W2005-00517-CCA-R3-CD) (raising sufficiency claim on grounds of disputed aggravating factor, but not DNA).

[100] 861 N.E.2d 898, 901 (Ohio Ct. App. 2006).

[101] 16 S.W.3d 156, 161, 172 (Tex. App. 2000).

[102] *Id.* at 167.

a sufficiency claim based on its conclusion that, simply because the RMP was 1 in 189 million for Hispanics, it was precluded from even considering the argument that "the DNA type *could* have been held by someone else in California or the world."[103] In *State v. Abdelmalik*, the court committed the fallacy in stating that "DNA evidence alone can provide sufficient evidence of identity to support a conviction" where DNA material found is inconsistent with casual contact and "there is a one in one quintillion likelihood that someone else was the source of the material."[104] Likewise, in *People v. Rush*, the court viewed a DNA match as "legally sufficient to support a guilty verdict" because, based on an RMP of 1 in 500 million, "the odds were 1 in 500 million that another person was the source of that DNA."[105] To date, no American trial or appellate court appears to have addressed the sufficiency of an uncorroborated DNA match based on a correct determination of the source probability.

In contrast to U.S. courts, British appellate courts appear to grasp that the determination of a DNA profile match's legal sufficiency depends on the source probability as determined by the RMP and the potential suspect population.[106] For example, in *Regina v. Lashley*, the Court of Appeal overturned a conviction based on an uncorroborated

---

[103] 35 Cal. Rptr. 2d 846, 859 n.27 (Ct. App. 1994), *aff'd*, 981 P.2d 958 (Cal. 1999). While the court did "take judicial notice" of the fact that there were only 30.4 million people in California at the time of the offense, *id.* at 851 n.13, an apparent sign that it understood the relevance of the number of potential suspects, it did not discuss why the potential suspect population would be limited to California residents, or why an RMP of 1 in 189 million among Hispanics would conclusively show that profile was unique in the relevant population.

[104] 273 S.W.3d 61, 66 (Mo. Ct. App. 2008).

[105] 630 N.Y.S.2d 631, 632, 634 (Sup. Ct. 1995). In at least one other case, a court did not explicitly commit the fallacy of the transposed conditional but appeared to reject a sufficiency claim with little analysis, and with no discussion of the RMP or the size of the suspect population. In *Springfield v. State*, 860 P.2d 435 (Wyo. 1993), a member of the Crow tribe was accused of a stranger rape in Sheridan, Wyoming. The court rejected a sufficiency claim without analysis based on the defendant fitting a general description of the perpetrator and a DNA match with an RMP of 1 in 221,000 for the American-Indian population and 1 in 17 million for the African-American population. *Id.* at 447–49. Springfield was three-fourths Crow and one-fourth African American. *Id.* at 437. Springfield was originally identified based on a lead related to the victim's missing camera rather than a cold hit, and the only evidence from the government at trial regarding the camera was exculpatory of Springfield. *Id.* at 437–38. While the victim testified at trial that Springfield "definitely resembles" her attacker, she could not identify him before trial, and her statement may merely have signified that he matched the description. *Id.* at 449. The court concluded that "the evidence was sufficient for reasonable and rational individuals to conclude that the appellant was the perpetrator." *Id.* at 449 (internal citation omitted).

[106] In British courts, as in the United States, the prosecutor bears the burden of proving sufficient evidence for every essential element of the crime. If the prosecution fails to meet this burden, the court should direct a verdict of acquittal. Adrian Keane, The Modern Law of Evidence 37 (7th ed. 2008).

cold hit, where the jury was told that seven to ten other male Britons likely shared the profile.[107] And in *Regina v. Watters*, the court held the evidence insufficient in a burglary case where the evidence against the defendant consisted of: 1) a DNA match with cigarette butts found at five burglary sites, 2) the fact that the defendant was a smoker and lived in the vicinity of the crimes, and 3) the fact that the defendant was male and most safe-crackers were men.[108] Crucial to the court's analysis was the fact that the defendant in this case had two brothers. While the RMP was 1 in 86 million assuming the defendant had no close relatives, the chance of a brother matching was 1 in 267, and neither of the defendant's two adult brothers had been ruled out as potential suspects.[109] The court ruled that the match statistics alone left too high a chance that one of the brothers may have committed the offense.[110]

Still, a British court ultimately upheld the defendant's conviction in a stranger-rape case, *Regina v. Adams*, in which the only evidence against the suspect was a DNA match with a disputed RMP somewhere between 1 in 2 million and 1 in 200 million and a potential suspect population of at least 200,000.[111] Moreover, while the victim initially described her attacker as being in his twenties, she later described Adams—upon viewing him—as being in his forties.[112] The defense argued that "the DNA evidence upon which the Crown had relied was incapable on its own of establishing guilt."[113] The court rejected this claim, concluding that "[t]here is, however, nothing inherent in the nature of DNA evidence which makes it inadmissible in itself or which justifies a special, unique rule, that evidence falling into such a category cannot found a conviction in the absence of other

---

[107] [2000] EWCA (Crim) 88, [3]; *see also* McCARTNEY, *supra* note 46, at 95 (discussing *Lashley*); Redmayne, *Rationality*, *supra* note 50, at 880 (asserting that RMP of defendant DNA match was around 1 in 4 million).

[108] [2000] EWCA (Crim) 89, [8], [10].

[109] *Id.* at [8], [9], [11], [12].

[110] *Id.* at [18]. Interestingly, the court later noted that even assuming the accuracy of a post-appeal recalculation putting the RMP for the brothers at 1 in 29,000, the evidence would still be insufficient: "[A]t the end of the day, greater though those odds are, they do nothing to eliminate the possible brother." *Id.* at [24]–[25].

[111] [1996] 2 Crim. App. 467, 468, 470; *see also* Kaye, *supra* note 6, at 484 (discussing *Adams*). The crime occurred in a town of just over 150,000 men between the ages of eighteen and sixty, and there seemed to be little dispute over treating the potential suspect population as at least 200,000, on the assumption that one could not state that there was more than a 75% chance that the perpetrator came from the town itself. *Adams*, [1996] 2 Crim. App. at 472–74; *see also* Kaye, *supra* note 6, at 484 (describing source population issues).

[112] *Adams*, [1996] 2 Crim. App. at 468; *see* Peter Donnelly, *Appealing Statistics*, SIGNIFICANCE, Mar. 2005, at 46, 46 (discussing *Adams*).

[113] *Adams*, [1996] 2 Crim. App. at 467.

evidence."[114] While the court may have been justified in observing that DNA evidence does not merit a "unique rule," the court never explained why the source probability in Adams's case established legally sufficient evidence of guilt.[115]

Since *Adams*, the Court of Appeal has reportedly developed guidelines for explaining to juries how to determine the source probability based on a DNA profile match, along the following lines:

> Suppose the match probability is 1 in 20 million. If you believe that number, then on average there will be 2 or 3 people in Britain whose DNA it could be, and probably no more than 6 or 7. . . . Now your job, as a member of the jury, is to decide, on the basis of the other evidence, whether or not you are satisfied that it is the person on trial who was the assailant, rather than one of the few other possible people who match. We don't know anything about the other people who match, although they are probably spread all over the UK, may have been nowhere near the scene of the crime, and some or all may also be ruled out by other factors, for example, gender or age.[116]

In sum, there appear to be three critical shortcomings in how courts—both in the United States and in the United Kingdom—have dealt with the issue of legal sufficiency in pure cold hit cases. First, while courts appear to understand that the sufficiency of a DNA profile match is a function of how high the source probability is, courts do not understand how to calculate the source probability, typically conflating it with the RMP and failing to consider the size of the suspect population. Second, although courts appear to have found some match statistics sufficient and others insufficient, no court has explained how it made this determination—that is, whether there is some numerical threshold that the source probability must meet to render the DNA match alone legally sufficient evidence of guilt. Finally, neither courts nor litigants appear to have explored the more fundamental question of whether numerical evidence alone, even a very high source probability, should ever be deemed sufficient evidence of guilt by itself.

One reason for courts' mathematical shortcomings is their traditional reluctance to view evidence in terms of an ultimate source probability calculated using Bayes' Theorem (such as a 91% chance

---

[114] *Id.* at 470.

[115] The appellate court quashed Adams's first conviction, based on concerns that the trial court had failed to properly instruct the jury. *Id.* at 482. In an appeal from Adams's subsequent second conviction on the same evidence, the court upheld the conviction. R v. Adams, [1997] EWCA (Crim) 2474.

[116] Donnelly, *supra* note 112, at 48 (noting that Court of Appeal now "advocate[s] that judges should summarise cases in the . . . way" described above).

the defendant is the source of the DNA) rather than a frequency statistic (such as the RMP, the frequency of a given profile in the population based on population tables).[117] There is nothing per se controversial about the universally accepted Bayes' Theorem, but courts have shunned its use in the courtroom out of fear that using mathematical formulae to solve the ultimate riddles of the case will either be confusing or involve ill-conceived attempts to quantify the probative value of impressionistic evidence.[118] The fear is understandable; obvious danger lurks in allowing expert witnesses to testify before the jury, for example, that a defendant's prior sexual assault conviction increases his odds of guilt in a rape case by fifteen percent. Even if such a valuation were not arbitrary, it surely invades the province of the jury.

But in a pure cold hit case, a trial court should have no reluctance in using Bayesian reasoning to determine legal sufficiency because the court would not have to arbitrarily attempt to quantify impressionistic evidence. Rather, it would simply determine the source probability, resolving disputes over the size of the relevant suspect population and RMP in favor of the government.[119] Because courts do not understand that the RMP is not equivalent to the source probability, they do not appreciate that the only way of determining the source probability is to use Bayesian reasoning, incorporating the RMP and potential suspect population. Given defendants' due process right not to be tried on legally insufficient evidence, courts must simply recognize that they have no choice but to calculate the source probability correctly.

---

[117] *See, e.g.*, State v. Skipper, 637 A.2d 1101, 1106, 1108 (Conn. 1994) (concluding that "[t]he utilization of Bayes' Theorem by the prosecution," specifically admission of evidence that prior probability of defendant's paternity of victim's child in sexual assault case was fifty percent, violated presumption of innocence and was reversible error); State v. Spann, 617 A.2d 247 (N.J. 1993) (reversing sexual assault conviction based on prejudicial admission of arbitrary fifty percent prior probability of paternity and, while not precluding potential discussion of prior probabilities on retrial, noting controversy surrounding "the evidentiary use of Bayes' Theorem at all," and Professor Kaye's suggestion that Bayes' Theorem be excluded from criminal cases).

[118] *See* Friedman, *supra* note 84, at 873–74 (noting "continuous battle" between frequentists, or "Bayesioskeptics," and Bayesians); F. Taroni & A. Biedermann, *Inadequacies of Posterior Probabilities for the Assessment of Scientific Evidence*, 4 LAW PROBABILITY & RISK 89, 92 (2005) (discussing concern that while "Bayesian view has secured its position as a coherent framework for evaluating evidence in forensic science," experts might usurp jury function by suggesting prior probabilities).

[119] Some disputes should be easy to resolve in the government's favor, such as in *Regina v. Adams*. *See supra* note 111 and accompanying text (discussing facts of *Adams* that court found sufficient to justify finding of proof beyond a reasonable doubt). Others might be more difficult and may preclude resolution of the case at the sufficiency stage. *See* discussion *infra* Part V (discussing such cases).

Courts' failure to address the other two questions—whether a source probability alone can ever be sufficient evidence of guilt and, if so, how high a source probability must be to attain legal sufficiency—may stem from their traditional hesitance to view standards of proof in quantifiable terms.[120] To decide whether a 91% source probability is legally sufficient, courts would presumably have to decide whether the reasonable doubt standard in a pure cold hit case is itself higher or lower than 91%. One typical concern with viewing "reasonable doubt" numerically is that "to attempt to quantify proof beyond a reasonable doubt changes the nature of the legal concept of 'beyond a reasonable doubt,' which seeks 'abiding conviction' or 'moral certainty' rather than statistical probability."[121] There are other ethical and conceptual concerns: that the jury should decide the standard in light of the facts of the case; that quantifying the standard too explicitly advertises the system's overt acceptance of erroneous convictions; and that "some probabilities may be inherently incapable of being given a precise number."[122]

But when the strength of the government's evidence has a quantifiable limit to begin with, the arguments for avoiding quantification of a standard of proof, at least for purposes of determining sufficiency, fall away. Unlike the speculative business of deciding whether an eyewitness's testimony is so persuasive as to constitute 95% certain proof of guilt, a judge can determine precisely whether a source probability of 91%—acting as an upper bound on the probability of guilt itself—

---

[120] *See, e.g.*, Commonwealth v. Sullivan, 482 N.E.2d 1198, 1200 (Mass. App. Ct. 1985) (holding that judge's charge requiring jurors to find only "above fifty percent" chance of guilt was error because "[t]he idea of reasonable doubt is not susceptible to quantification; it is inherently qualitative"); McCullough v. State, 657 P.2d 1157, 1159 (Nev. 1983) (holding that judge's charge equating reasonable doubt with 7.5 on scale from 0 to 10 was reversible error and noting that "[a]ny attempt to quantify [the concept of reasonable doubt] may impermissibly lower the prosecution's burden of proof, and is likely to confuse rather than clarify"); Rita James Simon & Linda Mahan, *Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom*, 5 LAW & SOC'Y REV. 319, 329 (1971) (quoting judge's view that "[p]ercentages or probabilities simply cannot encompass all the factors, tangible and intangible, in determining guilt—evidence cannot be evaluated in such terms"). *But see* Jack B. Weinstein & Ian Dewsbury, *Comment on the Meaning of "Proof Beyond a Reasonable Doubt,"* 5 LAW PROBABILITY & RISK 167, 172–73 (2006) (supporting adoption of jury instruction that would suggest possible quantification of proof beyond reasonable doubt as ninety-five percent).

[121] Commonwealth v. Rosa, 661 N.E.2d 56, 63 (Mass. 1996).

[122] James Franklin, *Case Comment—United States v. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): Quantification of the "Proof Beyond Reasonable Doubt" Standard*, 5 LAW PROBABILITY & RISK 159, 159–60 (2006) (listing objections to quantification of reasonable doubt as prologue to thesis that courts should at least set floor, such as 80%, below which evidence should be deemed insufficient); *see also* 9 JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 2497 (James H. Chadbourn ed., rev. ed. 1981) ("[N]o one has yet invented or discovered a mode of measurement for the intensity of human belief.").

meets that threshold. And while it may be difficult to agree on the precise moment where the numbers are compelling enough to go to a jury, there should be universal agreement in at least some cases. For example, if the upper bound on the probability of guilt is 49%, any judge should presumably determine that the evidence is insufficient as a matter of law, given that the "reasonable doubt" standard is at least higher than the greater-than-fifty-percent preponderance-of-the-evidence standard applicable in civil cases.[123] Thus, any difficulty in quantifying burdens of proof or evidentiary value at the margins should not stop courts from setting a floor below which pure cold hit evidence is agreed to be insufficient.[124]

The issues remaining to be resolved with respect to sufficiency in a pure cold hit case are thus: (1) whether there is some valid theoretical objection, improvidently ignored by litigants and courts, to allowing a conviction based solely on a numerical source probability; and (2) if not, how courts should determine whether a source probability in a particular case is compelling enough to constitute proof beyond a reasonable doubt as a matter of law.

III
IS A DNA MATCH STATISTIC ALONE INHERENTLY INCAPABLE OF JUSTIFYING A CRIMINAL CONVICTION?

In this Part, I consider an acknowledged but largely unexplored potential objection to a conviction based solely on DNA source probabilities: that naked statistical evidence of guilt is incapable of inspiring a juror's actual belief in guilt, which is more than a mere acknowledgment of a high likelihood of guilt—and therefore cannot inspire the "moral certainty" underlying the modern reasonable doubt standard. The question whether probabilistic evidence alone can justify a criminal conviction has been debated in the legal literature for four decades. Yet pure cold hit cases, many of which may involve RMPs with denominators in the sextillions—and thus astronomically high source probabilities regardless of the size of the suspect population—arguably eclipse the theoretical concerns raised by trial-by-mathematics scholars. As I explain, when source probabilities are high enough, they are effectively transformed into statements of certainty

---

[123] *See* 32A C.J.S. *Evidence* § 1627 (2008) ("The preponderance-of-the-evidence standard of proof requires that the factfinder determine whether a fact sought to be proved is more probable than not . . . .").

[124] *See* Franklin, *supra* note 122, at 159 (noting objections to quantifying reasonable doubt but advocating floor of 80% below which jurors should be instructed to acquit); *see also* discussion *infra* Part IV (proposing 99.9% source probability threshold before case is sent to jury).

rather than of probability. Though still fallible, such high source probabilities are capable of inspiring moral certainty.

I first explain why the "reasonable doubt" standard requires factfinders to reach an "actual belief" in, rather than merely an acknowledgment of a high probability of, the defendant's guilt. I next explain, by invoking trial-by-mathematics literature, why evidence perceived by factfinders as purely probabilistic does not meet this standard. Finally, drawing inferences from empirical studies on juror psychology, I explain why astronomically high source probabilities can inspire moral certainty because they are experienced by factfinders as assertions of certainty rather than probability.

A. Why the Government's Evidence Must Inspire an "Actual Belief" in Guilt To Meet the "Reasonable Doubt" Standard

A historical examination of the "reasonable doubt" test reveals that the certainty required to justify conviction in a criminal case—"moral certainty"—falls short of the metaphysical certainty of absolute "mathematical" or "demonstrative" proof of guilt but still requires that jurors reach an "actual belief" in the defendant's guilt. Thus, the government's evidence must inspire an actual belief in the defendant's guilt to be legally sufficient.

In the early days of the jury trial, jurors would be chosen based on their familiarity with the community, the parties, and the crime, often deciding cases based on their own knowledge of the events and character of those involved.[125] As society became more mobile and expressed a preference for juror impartiality, trials relied more on testimony and other secondhand evidence, and judges began to instruct juries on how to assess evidentiary value.[126] These developments coincided with the rise of humanism and the revival of skepticism in the sixteenth century, which presented potential threats to religious and scientific supremacy.[127] English theologians and naturalists met these challenges by recognizing an "intermediate level of knowledge"[128] that would justify belief and action even in the absence of "absolute"

---

[125] *See, e.g.*, Laura I. Appleman, *The Lost Meaning of the Jury Trial Right*, 84 IND. L.J. 397, 406 (2009) (noting juries were selected from "the immediate neighborhood" and were "chosen for their knowledge of the crime or their ability to find out").

[126] *See* Barbara J. Shapiro, *"To a Moral Certainty": Theories of Knowledge and Anglo-American Juries 1600–1850*, 38 HASTINGS L.J. 153, 155–56 (1986) (discussing shift toward juries as third parties and introduction of credibility standards); *see also* Appleman, *supra* note 125, at 406 (noting that juries began to lose their ability to personally investigate cases when society grew complex).

[127] Shapiro, *supra* note 126, at 156–57.

[128] *Id.*

or "metaphysical certainty."[129] These scholars distinguished between "mathematical" evidence "established by logical demonstration such as the proofs in geometry" and "moral" evidence, the ambiguous, impressionistic evidence "based on testimony and secondhand reports of sense data," that jurors actually encountered in real cases.[130] While unable to promise "mathematical certainty,"[131] moral evidence could promise "moral certainty," the highest level of certainty "possible from human, and necessarily fallible, sources."[132]

The moral certainty standard evolved into a requirement that the events alleged by the government be "so certain as not to admit of any *reasonable doubt* concerning them."[133] As "reasonable doubt" instructions became commonplace in the 1800s, courts continued to use the "moral certainty" language to explain the level of belief required to convict.[134] One nineteenth-century scholar explained that to "justify the inference of guilt" under a moral certainty standard, "the inculpatory facts must be incompatible with the innocence of the accused, and incapable of explanation upon any other reasonable hypothesis than that of his guilt."[135] In addition, nineteenth-century treatises recognized an "ethical significance of the juror's role" stemming from the

---

[129] Elisabeth Stoffelmayr & Shari Seidman Diamond, *The Conflict Between Precision and Flexibility in Explaining "Beyond a Reasonable Doubt,"* 6 PSYCHOL. PUB. POL'Y & L. 769, 770 (2000) (discussing origin of "moral certainty" as strategic counterpoint to intellectual challenges to "absolute certainty" and "metaphysical certainty").

[130] Shapiro, *supra* note 126, at 157–58.

[131] ARTHUR P. WILL, A TREATISE ON THE LAW OF CIRCUMSTANTIAL EVIDENCE 303 (Philadelphia, T. & J.W. Johnson & Co. 1896) (contrasting "mathematical" with "moral certainty").

[132] Robert C. Power, *Reasonable and Other Doubts: The Problem of Jury Instructions*, 67 TENN. L. REV. 45, 65 (1999); *see also* Victor v. Nebraska, 511 U.S. 1, 10–11 (1994) ("Moral evidence has for its subject the real but contingent truths and connections, which take place among things actually existing. . . . With regard to moral evidence, there is . . . real evidence on both sides. On both sides, contrary presumptions, contrary testimonies, contrary experiences must be balanced." (quoting 1 WORKS OF JAMES WILSON 518–19 (James DeWitt Andrews ed., Chicago, Callaghan & Co. 1896))).

[133] Shapiro, *supra* note 126, at 158 (emphasis added) (quoting JOHN WILKINS, OF THE PRINCIPLES AND DUTIES OF NATURAL RELIGION 8 (London 1675)).

[134] Consider, for example, the instruction in *Commonwealth v. Webster*:
> [E]very thing relating to human affairs, and depending on moral evidence, is open to some possible or imaginary doubt. It is that state of the case, which . . . leaves the minds of jurors in that condition that they cannot say they feel an abiding conviction, to a moral certainty, of the truth of the charge. . . . [T]he evidence must establish the truth of the fact to a reasonable and moral certainty; a certainty that convinces and directs the understanding, and satisfies the reason and judgment . . . . This we take to be proof beyond reasonable doubt . . . .

59 Mass. (5 Cush.) 295, 320 (1850). *See generally Victor*, 511 U.S. at 11–12 (describing moral certainty as "equivalent" to reasonable doubt and discussing history of concepts).

[135] WILLIAM WILLS, AN ESSAY ON THE PRINCIPLES OF CIRCUMSTANTIAL EVIDENCE 262 (Sir Alfred Wills ed., 5th ed. 1905).

"ethical notion of moral certainty."[136] This ethical significance required jurors to base their decisions not merely on reason, but on a "belief in the truth of events,"[137] such that the jurors are "satisfied in both 'understanding and conscience' that the evidence is sufficient to find guilt."[138] Others have similarly equated moral certainty with a "personal belief" in guilt.[139]

Judging from this historical account, the "moral certainty" standard might appear as merely a watered-down version of absolute certainty, with which mere mortals must make do given the unavailability of "metaphysical certainty" of guilt.[140] But by encompassing the "ethical notion" that jurors should not convict absent a personal "belief" in the guilt of the accused, the moral certainty standard is in a sense even more rigorous than a standard requiring mere acknowledgment of a very high likelihood of guilt.

---

[136] Steve Sheppard, *The Metamorphoses of Reasonable Doubt: How Changes in the Burden of Proof Have Weakened the Presumption of Innocence*, 78 NOTRE DAME L. REV. 1165, 1197 (2003) (discussing THOMAS STARKIE, A PRACTICAL TREATISE OF THE LAW OF EVIDENCE, AND DIGEST OF PROOFS IN CIVIL AND CRIMINAL PROCEEDINGS 513 (Benjamin Gerhard ed., Philadelphia, T. & J.W. Johnson, Law Booksellers 1842)). To be sure, philosophers "dubbed this sort of certainty 'moral' not because it had anything to do with ethics or morality but to contrast it with 'mathematical' certainty of the sort traditionally associated with rigorous demonstration." LARRY LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 33 (2006). When I describe a juror who lacks moral certainty I do not mean one who feels the evidence or prosecution itself is immoral or unethical, though a juror could feel that way and nullify. Rather, I mean a juror who is not convinced by the evidence to the point that he has an actual belief, in accord with his understanding and conscience, in the guilt of the accused.

[137] WILLS, *supra* note 135, at 6 (defining such "belief" as underlying concept of "moral certainty").

[138] Sheppard, *supra* note 136, at 1197 (quoting THOMAS STARKIE, A PRACTICAL TREATISE OF THE LAW OF EVIDENCE, AND DIGEST OF PROOFS IN CIVIL AND CRIMINAL PROCEEDINGS 477 (Benjamin Gerhard ed., Philadelphia, T. & J.W. Johnson, Law Booksellers 1842)).

[139] *See, e.g.*, Petitioner's Brief on the Merits at 25, Sandoval v. California, *consolidated with* Victor v. Nebraska, 511 U.S. 1 (1994) (No. 92-9049) (noting that Webster's defines "moral certainty" as "based on an inner conviction" (citing WEBSTER'S THIRD INTERNATIONAL DICTIONARY 1468 (1986))).

[140] This is certainly the manner in which the moral certainty requirement is characterized in the existing literature. *See, e.g.*, Power, *supra* note 132, at 65 (noting philosophers after 1600 recognized that "absolute knowledge based on moral evidence was impossible" and developed concept of moral certainty to "provide[ ] the highest degree of conviction possible from human . . . sources"); Shapiro, *supra* note 126, at 158 (same). Ironically, "moral certainty" has become the focus of litigation by defendants who argue that the phrase causes jurors to convict, in spite of a weak case, because of the moral gravity of the charged offense. *See, e.g.*, Victor v. Nebraska, 511 U.S. 1, 13–14 (1994) (discussing defendant's argument "that the phrase 'moral certainty' has lost its historical meaning, and that a modern jury would understand it to allow conviction on proof that does not meet the beyond a reasonable doubt standard").

B. Why Evidence Perceived as Probabilistic Fails To Inspire an "Actual Belief" in the Defendant's Guilt

In this Section, I invoke trial-by-mathematics literature and recent studies in cognitive psychology to explore whether a DNA match statistic alone, which is purely numerical but may establish a high subjective probability of guilt in jurors' minds, nonetheless might not inspire the "actual belief" in guilt needed to reach moral certainty.

A DNA match statistic is a probability. As such, it—like all modern scientific evidence—falls short of the historical ideal of metaphysically certain "mathematical evidence":

> The idea that science produced absolute truth, in contrast to the imperfect truths produced by non-scientific endeavors like law, was prevalent in the 18th century. That notion is no longer current among either scientists or philosophers of science because probability theory now holds sway. Today, all scientific knowledge is conceived as inherently probabilistic, and both scientists and philosophers of science would dispute the notion that science is characterized by the production of absolute certainty or truth.[141]

That such probabilistic evidence cannot inspire absolute certainty is, of course, not fatal as a matter of legal sufficiency so long as the evidence is capable of inspiring the "moral certainty" required by the reasonable doubt standard.[142] But is it? Some nineteenth-century jury instructions suggest no, telling jurors that, for example, "no degree of probability merely will authorize a conviction; but the evidence must be of such a character and tendency as to produce a moral certainty of the prisoner's guilt to the exclusion of reasonable doubt."[143] Why would even a high probability of guilt be unable to inspire an actual belief in guilt and, thus, moral certainty?

In certain respects, purely mathematical evidence such as a DNA match statistic would seem to be the least objectionable evidence one could have in a high-stakes case. After all, scientific evidence, as attorneys often point out to juries when disparaging alleged eyewitnesses,

---

[141] Simon A. Cole & Rachel Dioso-Villa, CSI *and Its Effects: Media, Juries, and the Burden of Proof*, 41 NEW ENG. L. REV. 435, 468 (2007); *see also* Susan Haack, *Irreconcilable Differences? The Troubled Marriage of Science and Law*, 72 LAW & CONTEMP. PROBS. 1, 12 (2009) (noting modern view that "scientific inquiry is by nature tentative and thoroughly fallibilist").

[142] While the phrase "moral certainty" has been removed from many modern jury instructions out of a concern that it might be misunderstood to the detriment of criminal defendants, the Supreme Court has recognized the phrase as synonymous with "beyond a reasonable doubt." *Victor*, 511 U.S. at 11–12 (describing these phrases as "equivalent").

[143] Sheppard, *supra* note 136, at 1210–11 (quoting State v. Jefferson, 10 So. 199, 200 (La. 1891)); *see also* State v. Gould, 395 So. 2d 647, 656–57 (La. 1980) (quoting and applying *Jefferson*); WILLS, *supra* note 135, at 6 (noting that belief "may be of various degrees, from moral certainty, the highest, to that of mere probability, the lowest").

is the type of evidence that doesn't lie, doesn't have an axe to grind, and doesn't forget. And notwithstanding the ubiquity of forensic science scandals in the past few years,[144] forensic identification techniques have played an important role in the exonerations of defendants who would otherwise have remained in prison based on notoriously unreliable nonscientific evidence.[145] Given that our society regularly sends people to prison on the basis of uncorroborated eyewitness testimony, which surely has a higher error rate than pure cold hit evidence, any issue with the latter cannot be one of verdict accuracy.[146]

And yet when faced with purely mathematical evidence of culpability, some courts and scholars feel a certain unease. In the famous "prison yard" hypothetical, for example, where a witness sees all but one of twenty-five prisoners approach and beat up a guard but cannot see who the one non-participant is, Charles Nesson has argued against convicting any particular prisoner, though the probability of any one prisoner's guilt is 24/25 or 96%, because he might be the one innocent.[147] And in the equally notorious "blue bus" hypothetical, where a plaintiff is struck and injured by a bus on a road at night, sees nothing

---

[144] *See, e.g.*, Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 5 & n.9 (2009) (discussing scandals involving faulty work from preeminent crime labs).

[145] *See* The Innocence Project, Facts on Post-Conviction DNA Exonerations, http://www.innocenceproject.org/Content/351.php (last visited July 26, 2010) (providing statistics on post-conviction DNA exonerations). Any arguments for caution in allowing a DNA match statistic to be the sole basis for conviction does not apply to the use of DNA to *exonerate* suspects. While interpretive errors in assuming number of contributors and testing errors such as allelic dropout could render a reported exclusion unreliable, an exclusion would not be challenged on the basis of the strength or weakness of the match statistic. Only a match needs a corresponding statistic to explain its probative value; an exclusion is simply the absence of a match. *See generally* Murphy, *supra* note 17, at 492–93 (discussing difference between using DNA as evidence of inculpation and evidence of exculpation).

[146] *See, e.g.*, Jonathan J. Koehler & Daniel N. Shaviro, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 247, 250 (1990) (suggesting that factfinder motivated solely by verdict accuracy would embrace overtly probabilistic methodology); Daniel Shaviro, Commentary, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 532–33 (1989) (suggesting that trial-by-mathematics critique values appearance of justice over verdict accuracy). Many jurisdictions prohibit convictions based on uncorroborated confessions, but this rule is related to the corpus delicti doctrine and is "designed to prevent the conviction of the coerced and mentally unstable for fictitious crimes," rather than concerned with the accuracy of the verdict. David A. Moran, *In Defense of the Corpus Delicti Rule*, 64 OHIO ST. L.J. 817, 817 (2003).

[147] *See, e.g.*, Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1192–93 (1979) (presenting prison yard hypothetical). L. Jonathan Cohen offers a similar hypothetical where 1000 people are seated at a rodeo but only 499 have paid. Cohen argues that "our intuitions of justice revolt against" imposing liability on any one patron for nonpayment, though the chance of nonpayment is

of the bus except that it is blue, and proffers evidence that the Blue Bus Company runs 80% of the blue buses along that route, some scholars see a problem in imposing liability, even though 80% far exceeds the preponderance standard applicable in civil cases.[148] Why would it be "that the high likelihood but starkly numerical case is thrown out of court while the cases based on self-serving testimony or additional circumstantial evidence will be put to the jury"?[149] In the real case upon which the bus hypothetical is based, the court's reason for directing a verdict for the defendant, notwithstanding that "mathematically the chances somewhat favor [the] proposition" that the defendant's bus struck the plaintiff, was that a rational prediction by the jury based on the evidence "was not enough" absent an "actual belief" in liability "in the mind or minds of the tribunal."[150]

Put differently, there exists a critical epistemological difference between, in the words of the nineteenth-century scholar William Wills, an actual "belief in the truth of events" and a mere "assent, which is the inevitable result of mathematical reasoning."[151] To analogize, imagine seeing a card flashed and, on the basis of the glimpse, believing the card is not a face card. If you were to learn later that the card is actually a king, you would presumably admit your belief was wrong. Now imagine blindly drawing a card from the deck and predicting the card not to be a king, given that only four out of fifty-two cards in the deck are kings. If you were to learn later that the card is actually a king, you would not think you had a mistaken belief; rather, you would merely conclude that your bet was wrong, and that "a relatively improbable event had occurred."[152]

---

greater than fifty percent for any given patron. L. Jonathan Cohen, *Subjective Probability and the Paradox of the Gatecrasher*, 1981 Ariz. St. L.J. 627, 627.

[148] While these are the essential facts of *Smith v. Rapid Transit, Inc.*, 58 N.E.2d 754 (Mass. 1945), the case has come to be referred to in the literature as the "blue bus hypothetical" and is commonly referenced in the trial-by-mathematics context. *See, e.g.*, Nesson, *supra* note 147, at 1194 (citing *Smith* in discussion of puzzling argument that, even in civil contexts, "high likelihood" but "starkly numerical" cases should be thrown out); Tribe, *supra* note 14, at 1340–41 (discussing *Smith* rule approvingly); *id.* at 1372–73 (describing imposition of liability based on purely probabilistic evidence as "intrinsically immoral").

[149] Nesson, *supra* note 147, at 1194.

[150] *Smith*, 58 N.E.2d at 755 (citation and internal quotation marks omitted).

[151] Wills, *supra* note 135, at 6.

[152] Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 Harv. L. Rev. 1357, 1361 (1985) (providing this example); *see also* Gary L. Wells, *Naked Statistical Evidence of Liability: Is Subjective Probability Enough?*, 62 J. Personality & Soc. Psychol. 739, 749 (1992) (noting Nesson's playing card example and fact that, in second scenario, person who guessed incorrectly based on statistics "was not to blame," having "made the correct decision," as "it was mere chance that produced the improbable outcome").

Why should we insist that jurors have an actual or personal belief in guilt, so long as they perceive a high subjective probability of guilt based on compelling probabilistic evidence? Why would the "ethical significance" of a juror's role preclude conviction based solely on the latter? One theory is that, if the public understands that jurors are merely betting on guilt rather than being personally convinced before condemning a potential innocent, it will perceive the system as under-valuing individual dignity:

> [T]here is a qualitative difference between the outcome of errone-ously convicting a man when the trier has been fully convinced of his guilt and . . . when the trier has reason to believe that he may be innocent. In the first of these situations, the trier is not called upon to make an explicit decision to risk morally condemning and pun-ishing an innocent man. In the second situation such a decision is required; it cannot be made without greatly undermining society's commitment to the dignity of the individual as an end in himself.[153]

When a juror forms an actual belief in the defendant's guilt, even if he knows that belief might be wrong, he "acquires an emotional stake" in the verdict, concludes that the event " 'really,' not just probably, hap-pened, and . . . forgets about the residual uncertainty" rather than "remaining acutely conscious of the possibility of verdict error."[154]

While a verdict based on qualitative evidence also carries with it a risk of convicting the innocent, a statistics-based verdict is unique in that the risk of error is overtly quantified. In the prison yard hypothet-ical, for example, we know that if we convict each of the 25 prisoners based on the 96% chance that any one of them is guilty, we are con-sciously sacrificing 1 innocent prisoner to convict the other 24. Con-victing a defendant "in the face of a recognized and quantitatively measured doubt in the particular case" risks public perception of injustice.[155] In contrast, "[a]s long as the concept [of reasonable doubt] is left ambiguous, members of the observing public may assume that they share with jury members common notions of the kinds and degree of doubt that are unacceptable."[156]

---

[153] Laurence H. Tribe, *An Ounce of Detention: Preventive Justice in the World of John Mitchell*, 56 Va. L. Rev. 371, 386 (1970).

[154] Shaviro, *supra* note 146, at 540.

[155] Tribe, *supra* note 14, at 1373.

[156] Nesson, *supra* note 147, at 1196; *see also* David Kaye, *The Laws of Probability and the Law of the Land*, 47 U. Chi. L. Rev. 34, 40 (1979) ("[W]e would prefer not to advertise the fact that we are willing to sacrifice one innocent person in order to secure the convic-tion of nineteen guilty ones."); Koehler & Shaviro, *supra* note 146, at 252 (observing that verdicts based on purely probabilistic evidence "make[ ] the risk of error explicit"); Shaviro, *supra* note 146, at 533 (noting objection based on "clarity of the risk of error").

The argument that statistical evidence alone fails, at least up to a point, to inspire an actual belief in guilt is supported by empirical work demonstrating some jurors' refusal to impose liability based on "naked statistical evidence" even when their subjective certitude[157] of culpability is high, and where they are willing to do so based on the same level of certitude in a case involving qualitative evidence, such as eyewitness testimony. In a groundbreaking study in 1992, psychologist Gary Wells gave mock jurors several versions of the "blue bus" hypothetical in which either the Blue or Grey Bus Company is responsible for running over a dog.[158] In one version, a weigh station attendant testified that a blue bus left the weigh station at 11:30 a.m., and the jurors heard that the accident occurred at 11:40 a.m. at a point about a ten-minute drive away from the station. The jurors also heard that the attendant was only correct 80% of the time, and that he mistook a blue bus for a grey one, or vice versa, the other 20%.[159] The other versions all involved either naked base-rate information (that the Blue Bus Company was responsible for 80% of the accidents along that road) or a case-specific statistic (that the tire tracks from the accident matched 80% of the Blue Bus Company's fleet of ten buses and only 20% of the Grey Bus Company's fleet of ten buses).[160] The jurors were more likely to impose liability based on the weigh station attendant's testimony than the versions with only naked probabilities.[161]

Subsequent studies have dubbed this phenomenon the "Wells Effect" and have attempted to isolate the reason for the effect.[162] One study tested a number of possible explanations, such as the theory that jurors prefer to believe human witnesses. In the end, the study validated what Wells believed was likely the case: Jurors "are more reluctant to convict when they can easily simulate a scenario in which the defendant is not guilty"; that is, "[w]hen probabilistic evidence of a defendant's guilt contains information that can be used to build a pos-

---

[157] By a "high subjective certitude" of guilt, I mean that the juror herself views the evidence as establishing a high probability of guilt. *See* Wells, *supra* note 152, at 741 (noting Laurence Tribe's different but related hypothesis that jurors might have only, say, 50% subjective certitude of culpability based on base-rate evidence suggesting 80% chance of culpability).

[158] *See id.* at 741–48 (describing experiment).

[159] *Id.* at 741.

[160] *Id.* at 742–43.

[161] *Id.* at 744.

[162] *See, e.g.*, Kevin Jon Heller, *The Cognitive Psychology of Circumstantial Evidence*, 105 MICH. L. REV. 241, 245 (2006) (discussing Wells Effect). *See generally* Keith E. Niedermeier et al., *Jurors' Use of Naked Statistical Evidence: Exploring Bases and Implications of the Wells Effect*, 76 J. PERSONALITY & SOC. PSYCHOL. 533 (1999) (describing subsequent testing of Wells Effect).

sible (even if unlikely) scenario in which another party is respon-
sible."[163] This "ease-of-simulation" theory[164] is consistent with the
"story model of juror decision making," according to which "jurors
explicitly construct alternative narrative accounts of the facts provided
at trial," assess the plausibility of those narratives, and choose the
most plausible story.[165] "[T]he easier it is to mentally simulate the
sequence of events in [a] story," the "greater its perceived plausibility
may be."[166] Because purely probabilistic evidence does not typically
make for a coherent narrative,[167] it may leave an alternative scenario
of innocence more psychologically plausible than would a confession
or eyewitness account with a similarly quantified error rate. On a sim-
ilar note, heuristics scholars have recognized a "certainty effect," or a
heightened sensitivity to even remote possibilities of innocence in
cases where the evidence of guilt is purely probabilistic.[168]

   To the extent pure cold hit evidence is perceived by jurors as
merely probabilistic, then, it is open to the criticism that it cannot
inspire an actual belief in the defendant's guilt. At most, it inspires a
belief that the evidence is strong and leaves the juror with an aware-
ness of an overt and quantifiable risk of convicting an innocent
person.

C.   Why Very High Source Probabilities Transcend the Realm of Probability and Become Assertions of Certainty Capable of Justifying Conviction

   While certain statistical evidence of guilt, such as in the prison
yard hypothetical, leaves the jury aware of the overt risk of convicting
the innocent, the match statistics in many pure cold hit cases are far
removed from the prison yard example. When "random match
probabilities approach[ ] a vanishing point,"[169] with denominators in
the trillions and beyond, they may well show an infinitesimally small
likelihood that any potential suspect other than the defendant would

---

   [163] Niedermeier et al., *supra* note 162, at 533, 541–42.

   [164] *Id.* at 533 (emphasis omitted).

   [165] *Id.* at 542 (discussing Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, in INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING 192–221 (Reid Hastie ed., 1993)).

   [166] *Id.*

   [167] *See, e.g.*, Heller, *supra* note 162, at 270 (noting that "abstract" probabilistic evidence "does not help [jurors] imagine how the defendant actually committed the crime"); RANDOLPH N. JONAKAIT, THE AMERICAN JURY SYSTEM 239 (2003) ("Jurors—and no doubt judges as well—often have difficulty integrating technical evidence into common-sense narratives. . . . This is particularly true of probabilistic information.").

   [168] *See, e.g.*, Heller, *supra* note 162, at 283 (explaining "certainty effect" whereby indi-viduals put less weight on merely probable outcomes).

   [169] LYNCH ET AL., *supra* note 70, at 345.

match the DNA profile.[170] If jurors experience such unfathomable numbers as "effectively impl[ying] certainty"[171] rather than mere probability, then a sufficiently high source probability may have the potential—like other fallible but absolute assertions about a defendant's guilt—to inspire an actual belief in, and thus moral certainty of, guilt.

Again, empirical work confirms such intuitions. Researchers have found that the Wells Effect begins to dissipate when the probability of the defendant's culpability reaches a high enough level.[172] In a follow-up article to the Wells study, researchers increased the probability that the Blue Bus company killed the dog to 99.9% and discovered that a majority—63%—of the participants were willing to find the company liable.[173] The authors suggested that the finding might reflect Wells's theory of "anticipated justification," according to which people "mentally simulate the possibility that the truth will be uncovered at a later time, and, if they are wrong in their verdict, they will need to justify it."[174] "Perhaps," the authors speculated, "an objective probability of 99.9% is so great that people feel able to justify their verdict."[175]

Alternatively, perhaps some of the participants experienced a 99.9% chance of culpability—or, equivalently, a 1 in 1000 chance of non-culpability—as beyond the realm of their experience and understanding, precluding their ability to visualize an alternative scenario of innocence. This view is in accord with "exemplar cueing theory": While juries tend to discount DNA match statistics when they can actually envision examples of other potential suspects in the population who might match, they will treat the match as "compelling proof" of guilt when they can no longer envision such examples.[176] "A DNA

---

[170] Again, the RMP is not equivalent to the source probability, and conflation of the two would be the so-called "fallacy of the transposed conditional." But if the RMP in a given case is lower than 1 in 6 trillion or so, the resulting estimate of the source probability—even assuming that the potential suspect population is the world's population of six billion—will exceed 99.9%. *See* discussion *infra* Part V (arguing for 99.9% source probability as possible starting point for legal sufficiency).

[171] LYNCH ET AL., *supra* note 70, at 345.

[172] *See* Edward F. Wright et al., *Factors Affecting the Use of Naked Statistical Evidence of Liability*, 136 J. SOC. PSYCHOL. 677, 685 (1996) (discussing study results supporting idea that high levels of objective probability moderate reluctance to rely upon naked statistics).

[173] *Id.*

[174] *Id.* at 685–86.

[175] *Id.* at 686; *see also* Heller, *supra* note 162, at 301 ("[A]lthough jurors are extremely sensitive to deviations away from certainty, research indicates they are generally willing to convict on the basis of probabilistic evidence that . . . establishes a 0.995 likelihood of the defendant's guilt.").

[176] Jonathan J. Koehler, *The Psychology of Numbers in the Courtroom: How To Make DNA-Match Statistics Seem Impressive or Insufficient*, 74 S. CAL. L. REV. 1275, 1280 (2001) (emphasis omitted).

match," then, "may make guilt overwhelmingly likely; the exculpatory inference(s) the defendant asks jurors to draw may stretch credulity to the breaking point."[177] As the probability of guilt becomes closer and closer to certainty along an asymptotic curve, a statistical statement of the likelihood of guilt may actually be transformed in jurors' minds from probabilistic evidence to individualized evidence justifying an actual belief in the defendant's guilt. One scholar, for example, in describing evidence in a hypothetical criminal case in which the defendant and perpetrator share a "freak genetic mutation which may be expected to occur no more often than once in 100 billion men"— making it "highly probable" that the defendant "is the only existing man" with the mutation—concluded that such a number was, in effect, "individualized evidence" of guilt.[178]

It may also be the case that near-certain statistical proof is more persuasive when the sacrificial innocent defendant is not known to exist. For example, some may still feel unease about convicting a defendant in the prison yard hypothetical, even if there were one million prisoners in the yard rather than twenty-five. However large the number, we still know that we are consciously sacrificing a known innocent defendant to convict the others. An analogy to popular culture would be the lottery—though the chances of picking the winning ticket are astronomically low, someone out there has a winning ticket. But in the blue bus case or a pure cold hit case, the only question is whether the defendant is culpable, and the chance of him not being culpable may well be infinitesimally small. So long as the defendant is truly culpable, there is no sacrificial innocent. The somewhat counter-intuitive notion that jurors might care more about the one sacrificial innocent prisoner out of a million than a much greater chance—say, 1 in 100,000—that the defendant on trial in a pure cold hit case might be wrongly accused based on a false coincidental match also may be partially explained by recent heuristics work exploring humans' preference for saving identified victims over preventing harm to unidentified potential victims.[179]

---

[177] Heller, *supra* note 162, at 283; *cf.* United States v. Veysey, 334 F.3d 600, 605 (7th Cir. 2003) (noting that in "obvious cases of fingerprint and DNA evidence," astronomically high match probabilities weaken "the case against allowing 'naked' statistical evidence to carry" burden of proof).

[178] Judith Jarvis Thomson, Rights, Restitution, and Risk: Essays in Moral Theory 247–48 (William Parent ed., 1986).

[179] *See, e.g.*, Karen E. Jenni & George Loewenstein, *Explaining the "Identifiable Victim Effect,"* 14 J. Risk & Uncertainty 235, 237–41 (1997) (reporting causes of "identifiable victim effect" where people are willing to expend greater resources to save identified lives than statistical lives). The "identifiable victim effect" would explain a juror's decision to acquit in the prison-yard example even if the yard had 10,000 prisoners (where the chance

"[U]nder certain conditions," then, DNA match evidence by itself may show that the "probability that anyone other than the accused in a criminal case was the source of [the evidence] is so low that the inference that the defendant was the source is 'virtually' or 'morally' certain."[180] The question becomes how these insights might inform a court's determination of legal sufficiency of the evidence in a pure cold hit case.

IV

CHOOSING A UNIFORM SUFFICIENCY THRESHOLD

A court determining legal sufficiency in a pure cold hit case could take one of three conceivable positions. First, it could take the position that an uncorroborated DNA profile match is always legally insufficient. Second, it could refuse to dispose of any pure cold hit case on sufficiency grounds. Or it could determine that some source probabilities are low enough to be legally insufficient absent corroboration, while other probabilities are high enough to send the case to the jury. In this Part, I explain why the first two positions are untenable and offer reasons for adopting a uniform sufficiency threshold below which courts should deem an uncorroborated DNA profile match legally insufficient. I then present possible arguments for choosing as a starting point a threshold of a 99.9% source probability.

The only justification for the first position—refusing to allow any conviction in a pure cold hit case—would presumably be the concerns raised in the trial-by-mathematics literature about the perceived inhumanity and illegitimacy of a system that bases conviction solely on mere assent to a high probability of guilt. As explained in Part III, however, many jurors would surely experience unfathomably high source probabilities as assertions of certainty, rather than probability, that—if credited—preclude the jurors from visualizing any possible scenario of innocence. If the source probability in a given case were high enough to transcend the bounds of probability and effectively become individualized evidence of guilt, then a rational juror might form an actual belief in guilt based on such evidence. The juror would

---

of any particular defendant being the one known innocent defendant is only 1 in 10,000), but to convict in a case in which there is no known innocent being falsely accused, but the chance of the defendant on trial potentially being falsely accused given the match evidence is 1 in 10,000.

[180] D. Michael Risinger, *The Irrelevance, and Central Relevance, of the Boundary Between Science and Non-science in the Evaluation of Expert Witness Reliability*, 52 VILL. L. REV. 679, 708 n.86 (2007); *cf.* Seth F. Kreimer & David Rudovsky, *Double Helix, Double Bind: Factual Innocence and Postconviction DNA Testing*, 151 U. PA. L. REV. 547, 609 (2002) (describing post-conviction DNA exoneration evidence as "provid[ing] morally certain proof" of innocence).

reach the moral certainty needed for conviction, and, while the verdict might still be inaccurate, the ethical principles underlying the moral certainty standard would remain intact.

The justification for the second position—sending all pure cold hit cases to the jury rather than declaring some cases legally insufficient—would presumably be based in deference to jury decision-making. Just as judges avoid second-guessing jurors who might view a seemingly incredible eyewitness as compelling, one might argue that judges should also defer to a hypothetical rational juror who might view a modestly strong source probability of, say, 80% as proof beyond a reasonable doubt. The reason judges defer to jurors' eyewitness credibility determinations, however, is not only that such determinations are inherently subjective, but also that jurors have a unique opportunity to assess the witness's demeanor. In contrast, verdicts based on evidence perceived as purely probabilistic involve factors unrelated to credibility that are equally knowable to those outside the jury[181]—such as the bus route statistics in the blue bus hypothetical—and that inspire only rational predictions based on the evidence. Unlike conflicting eyewitness testimony, a "coldly statistical" government case gives jurors "no opportunity to exercise" their skills of "perception or intuition," the very qualities that justify trial by jury to begin with.[182] Thus, a judge's ruling in a pure cold hit case that the evidence is insufficient would not require encroachment upon the jury's unique role as arbiter of credibility. At most, such a ruling would be in conflict with some jurors' own senses of when a source probability becomes an assertion of certainty.

While individual jurors might have different certainty thresholds, there are compelling reasons to adopt a uniform threshold for source probabilities below which a pure cold hit case should not go to the jury.[183] The first is that some source probabilities are low enough that the legal community should be able to agree that any rational juror—

---

181 *See, e.g.*, Nesson, *supra* note 152, at 1379 ("[G]iven the evidence's statistical nature, the public need not defer to the conclusion on the ground that the factfinder is in a better position to evaluate a witness's demeanor."); Nesson, *supra* note 147, at 1196 (arguing that "nothing presented to the jury puts it in any better position to judge" purely probabilistic cases, thereby decreasing confidence in need for, and efficacy of, jury in such cases).

182 Nesson, *supra* note 147, at 1196; *see also* Eleanor Swift, *Abolishing the Hearsay Rule*, 75 CAL. L. REV. 495, 504 (1987) ("One explanation of cases like *Smith* [*v. Rapid Transit*] is that only particularistic proof justifies the use of the elaborate system of adjudicative factfinding. Particularistic proof permits decisions to be based on the trier's knowledge and experience, rather than on the indiscriminate application of general probabilities.").

183 Note that the constitutional standard for determining legal sufficiency is the same in state and federal criminal trials. *See Jackson v. Virginia*, 443 U.S. 307, 315–16 (1979) (establishing legal sufficiency standard as due process standard under Fourteenth Amendment); *see also supra* notes 75–76 and accompanying text.

absent corroborative evidence—can envision a scenario of innocence. Obviously, a source probability of only 49%, lower than even the civil standard of proof, is insufficient. But even with a source probability of 90%, leaving a 10% chance that the defendant is not the source, it is difficult to fathom that a rational factfinder could not envision a scenario in which the defendant is innocent, absent any other evidence of guilt. Indeed, based on the Wells Effect studies, we have every reason to believe the contrary. Assuming the legal community comes to agree that certain source probabilities in cold hit cases are too low to form the basis for a rational juror's actual belief of guilt, judges should be bound to grant motions for acquittal in cases that do not meet a minimum threshold.[184] From a due process standpoint, there seems to be no reason to treat insufficient quantifiable evidence (such as a 90% source probability) differently from insufficient qualitative evidence (such as when the instrument used in a crime is not one deemed a deadly weapon).[185] Indeed, for a case—like pure cold hit cases—in which the evidence is inherently quantitative and thus more easily compared to the evidence in other cases of the same type, the need for a uniform, court-enforced sufficiency threshold is particularly strong.[186] Legal communities should thus use court rules, statutes, or the like to set numerical thresholds rather than allow individual trial judges to apply their own chosen thresholds.[187]

Another reason for imposing a uniform sufficiency threshold is to avoid perceptions of injustice or systemic illegitimacy in cases involving purely statistical evidence. Assuming the plausibility of the view that convictions based on purely numerical evidence undermine public confidence in verdicts because they are seen as an affront to human dignity, there is good reason to choose a sufficiency threshold that is both uniform and reflective of what we know about most

---

[184] *Cf.* Franklin, *supra* note 122, at 165 (noting that difficulties in deciding upon particular numerical threshold for reasonable doubt should not preclude legal community from imposing floor of 80%).

[185] *See supra* note 82 (citing case in which pellet gun was held not to meet definition of deadly weapon).

[186] *See, e.g.*, Craig R. Callen, *Cognitive Science and the Sufficiency of "Sufficiency of the Evidence" Tests*, 65 TUL. L. REV. 1113, 1121 (1991) ("[C]ourts also have an interest in using directed verdicts . . . as tools to promote uniformity or predictability in the enforcement of legal rules.").

[187] Presumably for this reason, judges in some jurisdictions direct a verdict of not guilty in operating-under-the-influence (OUI) cases where the breathalyzer result is close to the legal limit because of the machines' margin of error. *See, e.g.*, Randy S. Chapman, *Successfully Defending an OUI Case, in* 2 MASSACHUSETTS BASIC PRACTICE MANUAL § 13.5.3(c) (3d ed. 2009 & Supp. 2010) (noting that some judges in Massachusetts "will allow a motion for a required finding of not guilty as to the per se theory" when breath test results are close to state's legal limit).

people's certainty thresholds. Put differently, there is a value in using "the symbols of trial procedure[ ] to express society's fundamental commitment to the protection of the defendant's rights as a person, as an end in himself."[188] If most members of the public perceive a source probability of, say, 99.9999% as equivalent to a statement of certainty rather than probability, then arguably they will view a verdict based on such a source probability as reflecting the jurors' acceptance of an assertion of certainty based on the credibility of those who made the assertion, rather than the jurors' mere assent to a likelihood of guilt that, while high, still allows most of us to visualize alternate scenarios of innocence.

In turn, we have strong theoretical and empirical reasons to believe that most people's certainty thresholds in a pure cold hit case would be above a 99.9% source probability. The Wells Effect studies discussed in Part III reveal that even in a civil case, most people will not reach an actual belief in culpability below a 99.9% chance of culpability or, equivalently, a 1 in 1000 chance that the defendant is not culpable. If the simulation in these studies were changed to a criminal case, where the standard of proof is higher and the stakes are not monetary but involve physical liberty or even life, the probability of guilt may well have to be even higher to overcome the Wells Effect. For example, one recent study of jurors' reactions to blood group evidence accompanied by source probabilities in a criminal case found that over a quarter of jurors believed that a mere 0.5% "calculated risk" that the defendant was not the source of the evidence was still too high to be acceptable.[189] Another juror study testing "exemplar cueing theory" in a DNA case revealed that several jurors were still

---

[188] Tribe, *supra* note 14, at 1374. Tribe further explains why systemic legitimacy in this context matters:

> Methods of proof that impose moral blame . . . on the basis of evidence that fails to penetrate or convince the untutored contemporary intuition threaten to make the legal system seem even more alien and inhuman than it already does to distressingly many. There is at stake not only the further weakening of the confidence of the parties and of their willingness to abide by the result, but also the further erosion of the public's sense that the law's fact-finding apparatus is functioning in a somewhat comprehensible way, on the basis of evidence that speaks, at least in general terms, to the larger community that the processes of adjudication must ultimately serve.

Id. at 1375–76. *See also* Nesson, *supra* note 152, at 1367 (arguing that public "acceptance" of verdicts is critical to "promot[ing] our assimilation of the behavioral message embodied in the underlying substantive rule").

[189] Eric Magnusson, *Incomprehension and Miscomprehension of Statistical Evidence: An Experimental Study*, 1993 LAW MED. & CRIM. JUST. CONF., AUSTL. INST. CRIMINOLOGY 3, http://www.aic.gov.au/events/aic%20upcoming%20events/1993/~/media/conferences/medicine/magnus.ashx; *see also* Heller, *supra* note 162, at 284 n.237 (noting significance of this result in Magnusson's study).

able to envision the possibility of a coincidental match when faced with RMPs of 1 in 1 million and 1 in 1 billion in mock trials involving a Houston robbery-murder with no other evidence of guilt.[190] Specifically, some jurors given the 1 in 1 billion RMP still arrived at only an 83% subjective probability that the defendant was the source of the DNA.[191]

Based on these juror studies, a logical starting point for a sufficiency threshold is for courts to grant motions for judgment of acquittal whenever the source probability is lower than 99.9%. Such a sufficiency threshold might meet the force of the appearance-of-justice concerns underlying the trial-by-mathematics literature. It may turn out that, with more discussion and study, the legal community will decide that the sufficiency threshold for pure cold hit cases and other statistics-based prosecutions should be even higher.[192] Perhaps jurors are even more wary of DNA evidence than other types of naked statistical evidence, based on implicit concerns about the possibility of laboratory error or government abuse.[193] Whether or not such concerns are justified, the legal community might consider them in determining a sufficiency threshold for pure cold hit cases to promote the legitimacy of verdicts and to allow the concepts of "actual belief" and "moral certainty" to retain their meaning in such cases.

A 99.9% source probability threshold also comports with the scientific community's threshold for determining when it is appropriate to report "with reasonable scientific certainty" that "a particular individual is the source of an evidentiary sample."[194] While many scien-

---

[190] *See* Koehler, *supra* note 176, at 1293–96 (discussing experiment). The author does not make clear whether the attorneys in the mock trial presented the jury with only the RMPs or attempted to suggest a source probability based on the RMP and the size of the likely suspect population. Given that Houston is a city of several million, the source probability even given an RMP of 1 in 1 billion may have been lower than 99.9%. It is surely more likely that many jurors conflated the RMP with the source probability and yet still believed an objectively high source probability, absent other evidence of guilt, was less than fully compelling.

[191] Koehler cited as possible explanations for this phenomenon the Wells Effect and a few jurors' curious belief that a low RMP suggested a low source probability. *Id.* at 1295–96.

[192] At least one scholar has called for fingerprint analysts to move away from the individualization model and testify—in line with DNA analysts—in probabilistic terms, based on population databases and match statistics. *See* Simon A. Cole, *Forensics Without Uniqueness, Conclusions Without Individualization: The New Epistemology of Forensic Identification*, 8 LAW PROBABILITY & RISK 233, 249–50 (2009).

[193] *See* Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 J. LEGAL STUD. 107, 126 (2007) ("With the lab error rate statistics in particular, perhaps the jurors intuited the limits of the data and wanted better and more appropriate information about this lab.").

[194] Budowle et al., *supra* note 25.

tists refuse on principle ever to make assertions of uniqueness in court,[195] others have suggested or used a 99.9% source probability or its equivalent as a cutoff for source attribution.[196] To be sure, the legal community cannot delegate its choice of sufficiency threshold to the scientific community.[197] As scientists are the first to acknowledge, determinations of when the chance of finding another matching profile is remote enough to assert identity with "absoluteness" is a determination of philosophy, law, and psychology, and "not a statistical concept."[198] But the scientific community's view that a source probability of less than 99.9% leaves a reasonable possibility that the defendant is not the source seems to be at least relevant to the legal community's determination of whether a rational juror should still be able to visualize an alternative scenario of innocence with a source probability of, say, 95%. Moreover, concerns over systemic legitimacy might arise if the legal community allowed a conviction in a case where another large and influential community takes the position that

---

[195] *See, e.g.*, NAT'L RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 92 (1992) ("Regardless of the calculated frequency, an expert should—given the relatively small number of loci used and the available population data—avoid assertions in court that a particular genotype is unique in the population."); Song et al., *supra* note 32, at 23 ("[T]he lack of certainty makes claims of uniqueness improper to make in the presentation of DNA evidence in court.").

[196] *See, e.g.*, Balding, *supra* note 36, at 260 ("One of our assumptions was a 99.9% criterion for uniqueness, chosen arbitrarily."); Luttman Testimony, *supra* note 34, at 3-79 to 3-80 (noting that FBI's uniqueness threshold is RMP equal to about one thousand times U.S. population).

[197] I use the term "scientific community" in the same respect courts use it in the *Frye* context; that is, those scientists qualified "with sufficient training and expertise to permit them to comprehend" the scientific matter at issue. Blackwell v. Wyeth, 971 A.2d 235, 252 (Md. 2009). Here, the scientific matter would be the statistics involved in estimating the probability that a profile is unique. The relevant community in a matter involving DNA match statistics would include those in the derivative sciences of statistics, population genetics, and molecular biology, and not merely—or even primarily—forensic scientists. *See, e.g.*, United States v. Porter, 618 A.2d 629, 634–35 (D.C. 1992) (rejecting limitation of scientific community to forensic scientists for purposes of *Frye* test).

[198] B.S. Weir, *Forensics*, *in* 2 HANDBOOK OF STATISTICAL GENETICS 848–49 (2d ed. 2003); *see also* NAT'L RESEARCH COUNCIL, *supra* note 45, at 33 ("The definition of uniqueness is outside our province."); Budowle et al., *supra* note 25 (acknowledging in FBI publication that "[t]he size of the population and the appropriate confidence level to use" in making claims of profile uniqueness are, at bottom, "policy decisions"); Michael J. Saks & Jonathan J. Koehler, *The Individualization Fallacy in Forensic Science Evidence*, 61 VAND. L. REV. 199, 205 (2008) ("The concept of 'individualization' . . . exists only in a metaphysical or rhetorical sense. It has no scientific validity . . . ."); *id.* at 218 n.94 ("Although the DNA typing model has much to offer the traditional forensic sciences, offering source identifications at trial for sufficiently low probabilities would not be an implication of the science but an evasion of it in the service of advocacy.").

there remains a reasonable possibility the defendant is not the source.[199]

The 99.9% source probability is also consistent with how courts have treated DNA source attribution questions in other contexts. In paternity cases, a court will often presume a defendant's paternity if DNA testing shows with either 99% or 99.9% certainty that the defendant is the father.[200] Given that the standard of proof in paternity cases is less than "reasonable doubt" and constitutionally need only be a preponderance of the evidence,[201] the level of certainty required to convict someone in a criminal trial based on DNA testing should be higher than the level of certainty required to presume paternity in paternity cases, for reasons of systemic legitimacy.

To some, a 99.9% threshold may seem too favorable to the defendant. After all, most scholars, judges, and jurors would place the numerical equivalent of "beyond a reasonable doubt" in a typical case involving qualitative evidence at closer to 95% certainty of guilt or lower.[202] But when the inculpatory evidence in a case is purely probabilistic, these numbers shift upwards, according to the Wells Effect. While qualitative assertions of certainty from human witnesses can

---

[199] This is especially true given that phrases like "reasonable degree of scientific certainty" are amorphous terms that some experts take to mean merely "more probable than not" and others take to mean, more stringently, that "there was 'no reasonable or practical possibility that someone' other than the defendant" could be responsible. Molly Gena, Comment, *Shaken Baby Syndrome: Medical Certainty Casts Doubt on Convictions*, 2007 WIS. L. REV. 701, 716 (citations omitted) (discussing standards used to analyze bite marks).

[200] *See, e.g.*, ALA. CODE § 26-17-505(a)(1) (LexisNexis 2009) (creating rebuttable presumption with test result indicating 99% probability of paternity); IND. CODE § 31-14-7-1(3) (West 2008) (setting presumption of paternity at 99% probability); LA. REV. STAT. ANN. § 40:34(E)(5) (Supp. 2009) (setting conclusive presumption at 99.9% probability for purposes of child support); MICH. COMP. LAWS ANN. § 712.11(3) (West Supp. 2010) (setting presumption of paternity at 99% probability); 23 PA. CONS. STAT. ANN. § 4343(c)(2) (West 2001) (setting rebuttable presumption at 99% probability); TENN. CODE ANN. § 24-7-112(b)(2)(C) (2010) (setting rebuttable presumption at 99% probability); UTAH CODE ANN. § 78B-15-505(1)(a), (2) (LexisNexis 2008) (creating presumption with test results indicating 99% probability of paternity, rebuttable only by contradictory genetic test); WYO. STAT. ANN. § 14-2-705(a)(i) (2009) (setting rebuttable presumption at 99% probability).

[201] *See* Rivera v. Minnich, 483 U.S. 574, 575 (1987) (upholding preponderance of evidence standard for paternity tests). Some states require "clear and convincing evidence" for a posthumous claim of paternity. *See, e.g.*, LA. CIV. CODE. ANN. art. 192 (2008) ("The mother shall prove by clear and convincing evidence both that her former husband is not the father and that her present husband is the father.").

[202] *See, e.g.*, Lawrence M. Solan, *Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt*, 78 TEX. L. REV. 105, 126–27 (1999) (noting that judges' responses to survey asking them to quantify reasonable doubt ranged from 80% to 95% and higher, and jurors' responses ranged from upwards of 61%); Weinstein & Dewsbury, *supra* note 120, at 171 ("Were the parties to agree, we would like to add to our charge: 'In my opinion, a probability of guilt of no less than 95% should be necessary to support a conviction.'").

inspire actual belief even if they carry a decent likelihood of unreliability, statistical evidence of guilt must pass the certainty threshold—which typically does not occur below a 99.9% source probability—before being assessed by jurors in the same way as other fallible, but non-probabilistic, assertions.

Since there are no directed verdicts of guilt, the disposal of a criminal case at the sufficiency stage works only in the defendant's favor. One might wonder, however, whether pure cold hit cases should be an exception. After all, once the source probability becomes astronomically high, the government could argue that no rational juror could *fail* to find that the defendant is the source. This reasoning breaks down under closer scrutiny. When the source probability is below the threshold, after resolving factual disputes in the government's favor, the defendant is entitled to judgment of acquittal because no facts later found by the jury would make the case stronger. But when the source probability is above the threshold, jurors must still decide at trial whether the source probability is actually as high as the government claims it is and assess any exculpatory evidence. As to the reliability of the source probability, the defense may present evidence that, for example, the RMP reflects inaccurate assumptions about the independence of loci, and is in fact much higher,[203] or is dwarfed by a false positive rate that, if known, would show the DNA evidence to be weak.[204] And when there exists a "reality disjuncture" between the DNA match—the "genetic witness"—and "highly plau-

---

[203] A Stanford mathematician has called currently reported RMPs "ludicrous" and "well beyond the bounds of reality." Keith Devlin, *Damned Lies*, Devlin's Angle, Mathematics Ass'n of Am. (Oct. 2006), http://www.maa.org/devlin/devlin_10_06.html. Court-ordered studies of the Arizona, Maryland, and Illinois DNA databases suggest that nine-loci matches or even thirteen-loci matches may not be as rare as claimed. Jason Felch & Maura Dolan, *How Reliable Is DNA in Identifying Suspects?*, L.A. Times, July 20, 2008, at A1. The FBI continues to resist such studies of the CODIS database. Berkeley researchers have called for more RMP studies upon determining that, assuming independence of the STR loci, the average chance of a coincidental attribution in a cold hit case is 1 in 3.4 million. Song et al., *supra* note 32, at 22; *see also* Murphy, *supra* note 43 (calling for increased researcher access to CODIS).

[204] A false positive rate—even as low as 1 in 10,000 or 100,000—will dwarf a typical RMP in any source probability calculation. *See* William C. Thompson et al., *How the Probability of a False Positive Affects the Value of DNA Evidence*, 48 J. Forensic Sci. 47, 52 (2003) (showing that with RMP of 1 in 1 billion and prior odds of guilt being 1 in 1000, false positive rate as low as 1 in 10,000 still yields posterior odds of only 10 to 1 in favor of suspect being source of DNA); *see also* Balding & Donnelly, *supra* note 84, at 286 ("When extremely small match probabilities are claimed, it seems naive, at best, to ignore the possibility of false-positive results through human error."); Sangero & Halpert, *supra* note 71, at 45 (arguing against convictions based on single piece of evidence because of possibility of laboratory error).

sible" exculpatory evidence, the jury must still decide which to believe.[205]

V
HOW THE SUFFICIENCY THRESHOLD WOULD WORK IN AN ACTUAL CASE

To this author's knowledge, no pure cold hit case has been resolved at trial at the sufficiency stage based on the claim that the source probability is too low. As explained in Part II, courts' treatments of the sufficiency question in pure cold hit cases have been tainted by their failure to understand the difference between the RMP and the source probability, their reluctance to use Bayesian reasoning, and their hesitation to think of standards of proof in numerical terms.

Only some pure cold hit cases would be resolvable at the sufficiency stage. A defendant would only be entitled to a directed verdict when the source probability is less than 99.9% after resolving all disputes over the RMP and population size in the government's favor, and when there is no admissible, qualitative evidence to corroborate guilt. For example, in *Regina v. Lashley*, the British robbery case in which the only evidence linking the suspect to the robbery was a DNA match and the prosecution acknowledged that seven to ten other potential suspects may have shared the profile, the case presumably should have been resolved at the sufficiency stage in the defendant's favor,[206] since the RMP was 1 in 4 million.[207] The same goes for *Regina v. Watters*, the pure cold hit burglary case where the court held the evidence insufficient on appeal because the defendant's brother was a potential suspect and had a 1 in 267 chance of matching the profile.[208]

Some cases that would otherwise be resolvable on sufficiency grounds may be clouded by factual disputes over population size,[209]

---

[205] Cole & Lynch, *supra* note 7, at 47; *see also McDaniel* Amici Brief, *supra* note 94, at 15 ("[I]f the other evidence points to someone else, a jury could reasonably conclude that the match . . . was just a coincidence."); Balding, *supra* note 36, at 258, 260 (noting that non-DNA exculpatory evidence may trump even very small RMPs).

[206] *See* [2000] EWCA (Crim) 88, [v], [3], [5], [16] (describing facts in case).

[207] *See* Redmayne, *Rationality*, *supra* note 50, at 880 (discussing *Lashley*).

[208] *See* [2000] EWCA (Crim) 89, [9], [25], [29] (discussing facts and holding retrial not necessary).

[209] In resolving factual disputes about the size of the suspect population, courts should be careful not to limit the population to a particular class, such as a particular race or ethnicity, merely because the defendant is a member of that class. *See, e.g.*, People v. Wilson, 136 P.3d 864, 868 (Cal. 2006) (confirming that RMP should not be based on race in absence of evidence that perpetrator was of that race); D.H. Kaye, *Logical Relevance: Problems with the Reference Population and DNA Mixtures in* People v. Pizarro, 3 LAW PROBABILITY & RISK 211, 211 (2004) ("One strangely persistent fallacy in the interpreta-

the accuracy of the RMP, or the existence of arguably corroborative details. While disputes over such issues will often be properly left to the jury, courts should consider whether certain disputes can be resolved at the sufficiency stage by viewing disputed quantities in the light most favorable to the government or by determining as a matter of law that certain "corroborative" evidence, such as a prior conviction or ambiguous behavior consistent with innocence, is either only marginally relevant or the product of confirmation bias.[210]

Take the *Puckett* case, for example. Recall that, other than the defendant's 1977 conviction for sexual assault and evidence that he had ties to San Francisco at the time, the only evidence against him was a cold hit with an RMP of 1 in 1.1 million. If a court were to treat the "entire male population of the San Francisco Bay area between certain ages" as the suspect population, then the "initial population of potential suspects" would be, judging from Census data, over 2 million.[211] Multiplying this number by the RMP yields an "estimate of the number of unrelated individuals within this population who share the DNA profile taken from the victim's body."[212] Multiplying the population by an RMP of 1 in 1.1 million equals about 2, meaning that "Puckett is not the only man in the region who would have the requi-

---

tion of DNA evidence is that the relevant ethnic or racial population in which to estimate a DNA profile frequency necessarily is that of the defendant."). *See generally* Jonathan Kahn, *Race, Genes, & Justice*, 74 Brook. L. Rev. 325 (2009) (discussing problems with use of race in DNA evidence analysis presentation). Indeed, in some pure cold hit cases, the defendant is of a different ethnicity than the described perpetrator. *See, e.g.*, People v. Soto, 981 P.2d 958, 961 (Cal. 1999) (comparing victim's original description of rapist to defendant).

[210] For example, in the Brandon Mayfield fingerprint case, the absence of records showing that Mayfield had left the country was initially deemed by FBI investigators as evidence of concealment rather than innocence. Simon Cole notes that the bias inherent in having already targeted Mayfield may have contributed to officials' interpretation of his travel history as inculpatory. *See* Cole & Lynch, *supra* note 7, at 49–50 (discussing Mayfield story); *see also* D. Michael Risinger, *Cases Involving the Reliability of Handwriting Identification Expertise Since the Decision in* Daubert, 43 Tulsa L. Rev. 477, 594 (2007) ("[I]t is not clear, but is to be expected, that [the handwriting expert] generated his conclusions in the shadow of knowledge of the DNA 'match' variable which undermines reliability quite dramatically.").

[211] Kaye, *supra* note 6, at 491.

[212] *Id.* Multiplying the size of the suspect population by the RMP is, in essence, a "much simplified version" of Bayes' Theorem. *Id.* at 492; *see also* Thompson, *supra* note 58, at 11 n.2 (noting that population multiplied by frequency provides simplified estimate of chance of finding at least one match). This calculation assumes that the prior probability that the defendant is the source is one over the number of individuals in the potential suspect population—a "uniform prior probability distribution over most of the male population in the geographic region." Kaye, *supra* note 6, at 492. While the assumption that "*a priori*, each individual in the population is equally likely to have committed the crime" is "obviously not true," Song et al., *supra* note 32, at 24, it is presumably justified absent evidence from the government about other potential suspects.

site DNA profile."[213] In fact, the source probability given the match alone would be under 50%.[214] Assuming these numbers, Puckett would be entitled to a directed verdict if the case were based solely on the match.

But Puckett's case arguably contained qualitative corroborative evidence of guilt: his prior conviction for sexual assault. While such general evidence does not connect Puckett to the charged crime, many jurisdictions—including California—allow admission of sex convictions for purposes of proving propensity to commit a sex crime.[215] Moreover, if the court were to allow the case to go forward, the jury could also potentially discover corroborative evidence of guilt by viewing Puckett's courtroom demeanor, to see if it was somehow indicative of a guilty conscience.[216] Just as courts and lawmakers have developed rules about the corroboration required for certain specialized evidence to be deemed sufficient proof of guilt,[217] legal communities must decide what, if any, general inculpatory evidence can save an otherwise insufficient pure cold hit case.[218]

Toomes is a case in which disputes over the suspect population and RMP might preclude resolution at the sufficiency stage. Although

---

[213] Kaye, *supra* note 6, at 491–92.

[214] If the prior odds of Puckett being the source were about 1 to 2 million, multiplying by the inverse of the RMP (1.1 million) would yield posterior odds of around 1 to 2, or only about a 33% source probability.

[215] *See, e.g.*, FED. R. EVID. 413(a) (providing that such evidence "may be considered for its bearing on any matter to which it is relevant"); CAL. EVID. CODE § 1108(a) (West 2009) (providing that evidence of commission of another sexual offense is "not inadmissible").

[216] Courts are split as to whether the jury may properly consider a non-testifying defendant's courtroom demeanor as evidence. *Compare* United States v. Schuler, 813 F.2d 978, 981 (9th Cir. 1987) (holding prosecutor's comment about non-testifying defendant's demeanor improper), *with* Commonwealth v. Smith, 444 N.E.2d 374, 380 (Mass. 1983) (allowing prosecutor to comment on defendant's smirking during trial).

[217] *See, e.g.*, U.S. CONST. art. III, § 3, cl. 1 (stating that conviction for treason requires two witnesses to overt act or defendant's confession in open court); Moran, *supra* note 146, at 817 (discussing corpus delicti rule prohibiting conviction based solely on uncorroborated defendant's confession).

[218] While I make no definitive normative claim in this Article as to whether evidence such as prior convictions and courtroom demeanor should be enough to send an otherwise "pure cold hit" case to the jury, one could imagine compelling arguments to the contrary. For example, a juror's evaluation of demeanor after learning of a cold hit match could be tainted by confirmatory bias. While the juror might otherwise see a defendant's ability to look directly at jurors as suggestive of innocence rather than guilt-induced evasiveness, a tainted juror might view the same behavior as menacing. And while an extremely similar prior sexual act might suggest a propensity to commit similar acts, the use of the mere existence of a prior conviction to suggest guilt would seem to rest on general recidivism statistics that themselves might invoke appearance-of-justice concerns. Moreover, at least one publicized false cold hit pointed toward a man who, while innocent of the charged crime, had a prior conviction for murder. *See supra* note 59 (explaining that lab contamination likely led to false hit).

the court wrote that "*the only* . . . evidence connecting the defendant [Toomes] to the victim's rape is the DNA results"—a profile match with an RMP of 1 in 5 billion for African Americans and 1 in 22 million for Caucasians[219]—the jury also heard that the perpetrator was tall and male, that the police officer on the case knew Toomes's family, and that the crime happened in a small town, Ripley, about fifty miles north of Memphis.[220] A number of factors determine whether the RMPs in the case are enough for the source probability to rise to 99.9%:[221] whether the potential suspect population is, say, men in the town of Ripley (fewer than 4000),[222] the Memphis metropolitan area (around 500,000),[223] or the states of Tennessee and Arkansas (around 5 million);[224] whether the court decided to take race into account in the RMP even though there was no evidence of the perpetrator's race;[225] and, if so, what percentage of the suspect populations are African-American or Caucasian.

Similarly, in *People v. Rush*, a stranger rape case from Brooklyn, the RMP was reported as 1 in 500 million, with the only other evidence admitted against the defendant being that an acquaintance of his had seen him in the "vicinity" of the crime three days earlier.[226] It is difficult to discern from the opinion alone the significance of this additional evidence. Was the defendant scoping out a particular

---

[219] State v. Toomes, 191 S.W.3d 122, 127, 129 (Tenn. Crim. App. 2005).

[220] *Id.* at 124, 126.

[221] For example, if the suspect population were 5 million and the RMP were somewhere between 1 in 5 billion (for African Americans) and 1 in 22 million (for Caucasians), given that the suspect population is not entirely African-American, the chance of another person matching, using the simplified form of Bayes' Theorem, would be some number greater than (5 million x (1/5 billion)) = 1 in 1000. If the suspect population were only 4000, however, then even if the RMP were close to 1 in 22 million, the source probability would be well over 99.9%.

[222] *See* Table DP-1: Profile of General Demographic Characteristics: 2000, Geographic Area: Ripley, TN, http://censtats.census.gov/data/TN/1604763340.pdf (last visited July 25, 2010).

[223] *See* Table DP-1: Profile of General Demographic Characteristics: 2000, Geographic Area: Memphis, TN, http://censtats.census.gov/data/TN/390474920.pdf (last visited July 25, 2010).

[224] *See* Tennessee QuickFacts from the U.S. Census Bureau, http://quickfacts.census.gov/qfd/states/47000.html (last visited July 25, 2010); Arkansas QuickFacts from the U.S. Census Bureau, http://quickfacts.census.gov/qfd/states/05000.html (last visited July 25, 2010).

[225] *See supra* note 209 (urging caution in courts' decisions to limit population to particular class simply because defendant is member of that class).

[226] 630 N.Y.S.2d 631, 631–32 (Sup. Ct. 1995). The court viewed the complainant's failure to identify Rush at trial, and her identification of a courtroom spectator instead, as exculpatory evidence. *Id.* at 631–32. But the complainant did identify Rush as her attacker in a pretrial identification procedure. If the court had engaged in a pretrial sufficiency analysis, this additional information would have taken the case out of the subset of pure cold hit cases.

address? Or merely seen in Brooklyn that day? Assuming the suspect population is prime-aged men living in Brooklyn, probably around 680,000, the source probability would be less than 99.9%.[227] But any dispute over Rush's connection to the area would presumably send the case to the jury.

In some cases, the evidence will be clearly sufficient. Take, for example, the Ohio rape case *State v. Hunter*, in which the court asserted that "*literally no other evidence linked appellant to the crime*" other than the DNA match with an RMP of 1 in 756 trillion.[228] Even if one assumed the suspect population were as large as all men in the world between ages 18 and 65, there would be a smaller than 1 in 1000 chance that Hunter was not the source, given the RMP and suspect population.[229] Indeed, given that the suspect population for a crime will never be over about 6 billion, any RMP of 6 trillion or more (three orders of magnitude larger than the Earth's population) would be sufficient to show that the chance of the defendant not being the source is 1 in 1000 or less. As explained in Part I, although current typing methods routinely generate RMPs with denominators in the sextillions, cases involving very small amounts of DNA, degraded samples, a complicated mixture, mitochondrial DNA, or related suspects, will continue to result in much larger RMPs.

In still other more complicated cases, the chance the defendant is not the source given the match evidence might be almost exactly 1 in 1000, making any undisputed and highly compelling exculpatory evidence—even if not capable of being quantified—arguably relevant at the sufficiency stage. Take, for example, *Regina v. Adams*, the British rape case in which the only evidence against the suspect was a DNA match with a disputed RMP somewhere between 1 in 2 million and 1 in 200 million and a suspect population of at least 200,000.[230]

---

[227] Using the simplified version of Bayes' Theorem, the chance of non-uniqueness would be about 680,000 x (1/500 million) = 1/735.

[228] 861 N.E.2d 898, 901 (Ohio Ct. App. 2006) (emphasis added). There was also evidence that the perpetrator had stubble and smelled of alcohol, but no evidence that Hunter fit this description. *Id.* at 899–900. If Hunter had fit the description, the case would presumably be removed from the "pure cold hit" category unless an upper bound on the number of potential suspects meeting the description could be quantified. Given the lack of evidence that Hunter fit the physical description of the suspect, if the DNA match statistics suggested other matching profiles in the suspect population, these additional details—which could be just as true of nearly all other males aged eighteen to sixty-five with a matching DNA profile—presumably would not stop a court from granting a motion for judgment of acquittal.

[229] Using the simplified version of Bayes' Theorem, and taking 3 billion as an upper bound on the number of possible male suspects in the world, the chance of non-uniqueness would be less than 3 billion x (1/756 trillion) = 1 in 252,000.

[230] [1996] 2 Crim. App. 467.

Assuming an RMP of 1 in 200 million, the chance of the defendant not being the source based on the match evidence would be right around 1 in 1000.[231] While the case is on the cusp of the sufficiency threshold, the fact that the victim undisputedly looked at Adams and said he was around forty, whereas she thought the perpetrator was in his twenties, should perhaps be enough to entitle Adams to judgment as a matter of law. Though such evidence would be difficult to quantify[232] and its probative value would typically be left to the jury to decide, any marginal effect such evidence would have, however slight, would operate to decrease the probability of guilt below the 99.9% source probability. Courts will have to decide whether exculpatory evidence can ever be considered at the sufficiency stage, when circumstances indicate that its probative value, even in the government's view, is non-zero.

Finally, a recurring and difficult issue with which courts will also have to grapple is the unknown, but clearly non-zero, chance of a false positive from laboratory or interpretive error. At least one defendant has argued in a pure cold hit case—with an otherwise astronomically high source probability—that the error rate, if the government calculated it, would dwarf the RMP and render the DNA evidence alone insufficient.[233] The argument is compelling: In a pure cold hit case where the suspect population is, say, 10 million, a false positive rate of even 1 in 100 million—even if the RMP were infinitesimally small— would render the evidence insufficient under a 1-in-1000 certainty threshold. [234] While not all errors lead to false positives,[235] and the

---

[231] Using the simplified version of Bayes' Theorem, the chance of non-uniqueness would be about 200,000 x (1/200 million) = 1 in 1000.

[232] Statistician Peter Donnelly, testifying for the defense, offered a possible quantification of the exculpatory evidence in the case and then calculated the posterior odds of guilt, assuming an RMP of 1 in 200 million, as being only 55 to 1, meaning that the posterior probability of the defendant being guilty based on the evidence was 55/56, or 98.2%. *See Adams*, 2 Crim. App. at 476–77 (citing and discussing Donnelly testimony); *see also* Kaye, *supra* note 6, at 484 (describing Donnelly's application of Bayes' Theorem). The Court of Appeal admonished the defense for using Bayesian reasoning, though it appears the court's true complaint was with the attempt to quantify inherently impressionistic evidence in a way that may have confused the jury. *See Adams*, [1996] 2 Crim. App. at 481 ("[T]he theorem's methodology requires, as we have described, that items of evidence be assessed separately according to their bearing on the accused's guilt . . . . That in our view is far too rigid an approach to evidence of the type that a jury characteristically has to assess . . . .").

[233] *See* Memorandum in Support of Defendant's Motion for Judgment of Acquittal at 1–4, State v. Derr, No. K04-930 (Md. Cir. Ct. June 29, 2006) (arguing that defendant must be acquitted because failure to present error rate to jury undermines sufficiency of DNA evidence); *see also* Sangero & Halpert, *supra* note 71, at 47 (arguing that single piece of evidence should never be sufficient to convict).

[234] *See* Sangero & Halpert, *supra* note 71, at 54–55 (presenting table of probabilities of guilt as calculated under Bayes' Theorem).

error rate will presumably be lower in a cold hit case than a confirmatory one, [236] the rate may well be higher than 1 in 100 million. But no laboratory conducts the type of frequent proficiency testing that would allow it to state an error rate, and it is difficult to imagine any laboratory doing so absent a strong incentive or court order.[237]

CONCLUSION

Given the astronomically high source probabilities in some DNA cases, it may seem "as though 'DNA' has transcended the mortal realm and moral certainty of criminal evidence and taken its place in a dream world of unassailable scientific evidence and mathematical certainty."[238] On the contrary, DNA match statistics begin as purely probabilistic evidence incapable of inspiring moral certainty, yet transcend that realm at a certain point to take their place among other assertions of certainty that, while fallible, are capable of inspiring an actual belief in, and thus moral certainty of, guilt. Society's recognition of match statistics' ability to fully convince jurors of guilt is just as important as its recognition that such statistics fall far short of absolute, metaphysical certainty.

This paper has posited and confirmed the former proposition—match statistics' ability, under certain conditions, to fully convince jurors of the defendant's guilt—and has offered courts a guide in determining when, in a case that turns entirely on a DNA match statistic, the numbers are compelling enough to send to the jury. I intend the 1-in-1000 benchmark merely as a starting point in the continuing discussion of how to give meaning to the reasonable doubt standard in the modern age of statistics and science-based prosecutions.

The question of the lay jury's role in the coming era of technologically advanced prosecutions remains to be answered. Some have warned that "[w]ith the advent and rapid advancement of DNA technology, there is a danger that juries will be viewed as virtually dispen-

---

[235] *See, e.g.*, Peter Gill & Amanda Kirkham, *Development of a Simulation Model To Assess the Impact of Contamination in Casework Using STRs*, 49 J. FORENSIC SCI. 485, 486 (2004) ("The most probable outcome of a contamination event is a false exclusion.").

[236] *See* Thompson et al., *supra* note 204, at 53 ("[T]he particular circumstances of database searches would seem to rule out, or at least greatly reduce, the likelihood of some types of errors, such as those arising from switching or cross-contaminating samples . . . . However, other . . . errors, such as those arising from misinterpretation of test results, might still produce false matches.").

[237] The 1996 National Research Council Report rejected calls for requiring the type of blind proficiency testing presumably needed to estimate a laboratory-specific error rate or for using an industry-wide estimated error rate, on grounds that both measures would be infeasible and would not be representative of the chance of error in any particular case. *See* NAT'L RESEARCH COUNCIL, *supra* note 45, at 85–87.

[238] LYNCH ET AL., *supra* note 70, at 340.

sable in cases in which there is overwhelming scientific evidence of guilt."[239] At first glance, pure cold hit cases seem to require the jury to perform only two tasks: number-crunching and deciding whether the source probability given the evidence is high enough to justify conviction. The first is not one that would seem to call for resolution by trial, much less trial by lay jury. Rote mathematical operations, if all the necessary quantities were at hand, could just as easily—and more accurately—be performed by a non-adjudicative board of scientists. The second task also does not involve traditional jury functions. Rather, it is a moral or political line-drawing that society could remove from the purview of the jury and relegate to other institutional actors, such as the legislature or the scientific community, arguably without upsetting the jury's familiar role as factfinder or as a check on unjust laws or prosecutions.

But as this Article has indicated, the mathematical evidence in even the "purest" cold hit case does not simply involve number-crunching. Even in a pure cold hit case, the "DNA evidence is meaningful only when it is embedded in stories that mention other evidence, possible suspects, and how the evidence itself was handled and interpreted."[240] And while there are important systemic reasons for disposing of certain cases on sufficiency grounds, the defendant ultimately cannot be convicted unless each juror actually reaches a state of moral certainty of, or actual belief in, the defendant's guilt. In ensuring fair trials in pure cold hit cases, courts will have to grapple with jurors' inability to understand statistics, and with the compounding concern that, because of DNA's near-mythical status in modern culture, some jurors may not understand that they can still acquit in the face of astronomically high source probabilities. Further challenges arise given the line between rational mistrust of certain futuristic evidence and fear-based nullification, on the one hand, and the line between being fully convinced and unduly mesmerized by scientific evidence, on the other. As DNA "becomes reified as a machinery of truth for determining guilt and innocence,"[241] the legal community must ensure that both courts and juries understand the types of certainties—and uncertainties—such evidence entails when it is the sole proof of guilt.

---

[239] Julie A. Seaman, *Black Boxes*, 58 EMORY L.J. 427, 474 (2008).

[240] LYNCH ET AL., *supra* note 70, at 191.

[241] *Id.* at 346.