JOURNAL OF
FORENSIC SCIENCES

Fingerprint error rate on close non-matches

Jonathan J. Koehler PhD[1] | Shiquan Liu PhD[2]

[1]Northwestern Pritzker School of Law, Chicago, IL, USA

[2]Institute of Evidence Law and Forensic Science, China University of Political Science and Law, Beijing, China

Correspondence
Jonathan J. Koehler, Northwestern Pritzker School of Law, Chicago, IL, USA.
Email: jay.koehler@northwestern.edu

Abstract

The accuracy of fingerprint identifications is critically important to the administration of criminal justice. Accuracy is challenging when two prints from different sources have many common features and few dissimilar features. Such print pairs, known as close non-matches (CNMs), are increasingly likely to arise as ever-growing databases are searched with greater frequency. In this study, 125 fingerprint agencies completed a mandatory proficiency test that included two pairs of CNMs. The false-positive error rates on the two CNMs were 15.9% (17 out of 107, 95% C.I.: 9.5%, 24.2%) and 28.1% (27 out of 96, 95% C.I.: 19.4%, 38.2%), respectively. These CNM error rates are (a) inconsistent with the popular notion that fingerprint evidence is nearly infallible, and (b) larger than error rates reported in leading fingerprint studies. We conclude that, when the risk of CNMs is high, the probative value of a reported fingerprint identification may be severely diminished due to an elevated false-positive error risk. We call for additional CNM research, including a replication and expansion of the present study using a representative selection of CNMs from database searches.

KEYWORDS
close non-match, error rate, false-positive, fingerprint, fingerprint accuracy, fingerprint identification, proficiency test

1 | INTRODUCTION

For more than a century, fingerprint evidence has been accepted by investigators, courts, jurors, and the general public as one of the strongest and most accurate forms of evidence available in legal cases. In a typical case, a fingerprint examiner compares a latent (unknown) print with one or more exemplar (known) prints and then offers one of three conclusions: identification (the prints share a common source), exclusion (the prints come from different sources), or inconclusive. A fingerprint examiner's conclusions are rarely challenged [1] and widely presumed to be accurate [2].

Reports from prestigious scientific panels suggest that the scientific basis for the trust people place in fingerprint evidence is weak [3,4]. The President's Council of Advisors on Science and Technology

(PCAST) stressed that the foundational validity of the standard fingerprint method (ACE-V) can only be established by empirical studies that examine the rate at which examiners err [5]. This report concluded that there are "only two properly designed studies of the foundational validity and accuracy of latent fingerprint analysis" ([5], p. 101). In Ulery et al. [6], 169 experienced volunteer fingerprint analysts examined approximately 100 latent-exemplar print pairs, about 30% of which were non-mated (i.e., from different sources). Participants judged each pair to be an individualization, an exclusion, or an inconclusive. Five examiners committed a total of six false-positive errors out of 3628 decisions (excluding inconclusives) on non-mated pairs. This result yielded a false-positive error rate of 0.2% (95% C.I. = 0.1%, 0.4%, exact binomial). (The authors reported a false-positive error rate of 0.1% ([6], p. 7735), but this computation included 455 inconclusives).

In Pacheco et al. [7], 109 experienced volunteer fingerprint analysts examined various pairings from 80 latent prints that were produced from 10 known sources. As in Ulery et al. [6], participants judged each pair to be an individualization, an exclusion, or an inconclusive.

Twenty-eight examiners committed a total of 42 false-positive errors out of 995 decisions (excluding Inconclusives) on non-mated pairs, yielding a false-positive error rate of 4.2% ([7], p. 53). However, 35 of the 42 false-positives were errors involving the wrong finger or palm but the right person ([7], p. 64-5). Because the source hypothesis of interest is nearly always "this person" rather than "this particular finger of this person," these errors are generally inconsequential (but see Ref. [8]). Excluding these 35 errors, the false-positive error rate was 7 out of 960 decisions or 0.7% (95% C.I. =0.3%, 1.5%; but see Refs. [8,9] for alternate calculations). These seven errors were committed on image pairs that were, on average, of moderate difficulty as rated by the participants. The examiners who committed errors varied widely in terms of their fingerprint training and experience. Taken together, the two fingerprint error rate studies that meet the guidelines outlined in the PCAST report [5] indicate a false-positive error rate for volunteer examiners of less than 1%.

2 | AUTOMATED FINGERPRINT DATABASES AND CLOSE NON-MATCHES (CNMS)

Though methodologically sound, it is not clear how applicable [6] and [7] are to very difficult cases in which two prints from different people have many common features and few discernible dissimilar features. Such print pairs are known as close non-matches (CNMs) [10–12]. The risk of encountering a CNM is heightened when large databases are searched for the source of a print [13]. These databases, which contain prints from millions of individuals [14] and are used by most police departments in the United States [15], are searched hundreds of thousands of times each year. The use of these databases, particularly large ones, may increase the risk of a false identification because they may contain hard-to-distinguish CNM prints [16]. Concern about false identifications from database-derived CNMs is not merely theoretical. The Brandon Mayfield case, detailed in [12], shows both that database CNMs exist and may fool even the best fingerprint examiners.

3 | CNM STUDIES TO DATE

There has been little empirical work on CNMs and the risk of false-positive error that they pose. The work that does exist has not always been published in peer-reviewed journals or has not been a focal point in the articles in which the relevant data appear.

Table 1 summarizes the available CNM data.

In 1995, the proficiency testing agency Collaborative Testing Services (CTS) unwittingly produced an early CNM study [17]. Laboratory participation was voluntary, and laboratories were free to use the test as a competency test, a training exercise, or in whatever manner they pleased. The test materials included seven bloody latent prints (arbitrarily labeled 5a-5 g) and exemplar prints from four suspects. Two of the seven latent prints (5c and 5f) did not belong to any of the four suspects and therefore should not have been identified by the examiners. However, those two latent prints shared some features with one of the suspect's prints because they were produced by that suspect's twin brother whose exemplar prints were not provided. Latent print 5f and exemplar prints from each of the twin brothers were included in Expert Working Group on Human Factors in Latent Print Analysis [4] as figure 1.4 on page 6. Six out of 109 (5.5%) and 29 out of 112 (25.9%) examiners falsely identified prints 5c and 5f, respectively.

Langenburg et al. [18] published test data from 176 U.S. latent print examiners, each of whom examined 12 print pairs using a web-based interface. Five of the print pairs were CNMs. The examiners falsely identified 23 of the 788 CNM trials (2.9%).

Neumann et al. [19] conducted a study using a web-based interface that provided test data from 146 U.S. examiners and trainees each of whom examined 15 print pairs. One pair (Trial 12) was arguably a CNM because the median number of common minutiae observed was high [10]. Trial 12 was falsely identified as a same source pair by 11 out of 84 examiners (13.1%) who offered a definitive conclusion.

Liu et al. [20] asked 40 Chinese fingerprint examiners to evaluate five "difficult" print pairs. One pair (Mark 4) was a CNM. Mark 4 was falsely identified by 3 out of the 27 examiners (11.1%) who provided an identification or exclusion conclusion. The examiners who falsely identified Mark 4 varied widely in experience (1-25 years).

The studies discussed above and summarized in Table 1 suggest that the false-positive error risk for fingerprint CNMs may be substantially higher than the very low error risks that are testified to in court [21,22], described by the government [23], and believed by judges [24] and the general public [2]. However, the experience, quality, and motivation of the volunteer examiners in these early CNM studies are not always clear.

4 | THE STUDY

The present CNM study focuses on the conclusions reached by experienced, high-level non-volunteer fingerprint examiners in the context of a national proficiency test. In 2014, S.L. was given

TABLE 1 False-positive error rates for CNM fingerprint pairs in previous studies

1 Pair – "5C" Collaborative Testing Services, Inc. [17]	1 Pair – "5F" Collaborative Testing Services, Inc. [17]	5 Pairs Langenburg et al. [18]	1 Pair – "Trial 4" Neumann et al. [19]	1 Pair – "Mark 4" Liu et al. [20]
$\frac{6}{109}$ = 5.5%	$\frac{29}{119}$ = 25.9%	$\frac{23}{788}$ = 2.9%	$\frac{11}{84}$ = 13.1%	$\frac{3}{27}$ = 11.1%

permission by national authorities to conduct a proficiency test using the PiAnoS4 software program. This software program, which was used in [18,19], was developed by researchers at the University of Lausanne in Switzerland. The software was translated by S.L.'s research team.

4.1 | Participants

One hundred and twenty-five nationally accredited fingerprint agencies, employing approximately 500 fingerprint examiners, participated in this proficiency test.

One hundred and six of those agencies were located in a single, large portion of the country and the other 19 agencies were a random selection from a set of elite national forensic science agencies. The test was reviewed and approved by a national forensic director prior to distribution.

Unlike the studies referenced above, participation in this test was mandatory at the level of agency. The researchers complied with all ethical and professional requirements set forth by the host country prior to, during, and after conducting this test, and agencies were instructed to follow their usual careful examination protocols. Although we are not in a position to know whether a given agency was any more or less careful when completing this test than they are in actual casework, motivation to perform well was presumably high because errors on the test could have serious professional consequences for both the agency and the examiners. Because the unit of participation was by agency rather than examiner, some tests were completed by a single examiner and some were completed by a collaborative group of high-level examiners within the agency. This feature is consistent with how fingerprint analyses are conducted in casework throughout the country in which the test was conducted. Here we note that there are only two major differences between the approach taken by fingerprint examiners in the host country and fingerprint examiners in the United States In the host country, examiners will call an identification if and when they observe 12 or more minutiae in agreement without observing any minutiae in disagreement. In the United States, a numerical point threshold is not used. Further, in the host country, examiners are discouraged from using the inconclusive category. Print pairs are generally identified or excluded.

4.2 | Latent images

Each agency was provided with 5 latent-exemplar print pairs. Two print pairs were non-mated CNMs that we refer to here as CNM_1 and CNM_2. CNM_1 was included in Neumann *et al* [19] as figures 42 and 43 on page 80, and as the Trial 12 figure on page 94. CNM_2 was included in Clark [25]. The remaining three print pairs were mated pairs from three different individuals. The agencies were not told which, if any, of the five print pairs were mated or non-mated, nor were they given any other information about the prints or cases from which the prints were obtained.

5 | METHOD

The study employed a web-based platform that used Picture Annotation Software 4 (PiAnoS4) fingerprint software. PiAnoS4 enables participants to conduct a detailed, annotated examination of print pairs using the ACE-V method. All agencies received a PiAnoS4 manual and detailed instructions for completing the test. One practice trial (a same source print pair) was provided at the outset to allow users to become familiar with the PiAnoS4 program. The software initially revealed the unknown (latent) print only. After the agency completed an analysis of this print, the software revealed the known (exemplar) print for comparison. Examiners were instructed to provide one of three conclusions for each latent-exemplar print pair: identification, exclusion, or inconclusive. In keeping with casework practices in this country, the instructions discouraged reliance on the inconclusive conclusion. Examiners were also encouraged to provide written comments that document their thoughts at various decision points.

Each agency received a unique user name and was given 5 days to complete the test and submit results. The test was designed to take 2-5 h to complete. Three agencies reported technical difficulties with the software when using the Internet Explorer browser. These problems were resolved by shifting to either the Chrome or Firefox browsers. Regular reminders to complete the test were sent, and all agencies completed with the test within the allowable 5-day time frame. Unlike proficiency test practices in many other countries, the agencies did not submit their results anonymously.

6 | RESULTS

All 125 agencies provided results (100% participation rate). Ninety-one percent of agencies indicated that CNM_1 had value for identification and 99% indicated that CNM_2 had value for identification. Eighteen agencies (14.4%) provided an inconclusive for CNM_1 and 29 agencies (23.2%) provided an inconclusive for CNM_2. Among the 107 agencies that reached a definitive conclusion for CNM_1, 17 erroneously reported an identification, yielding a false-positive error rate (excluding inconclusives) of 15.9% (95% C.I.: 9.5%, 24.2%). Among the 96 agencies that reached a definitive conclusion for CNM_2, 27 erroneously reported an identification, yielding a false-positive error rate of 28.1% (95% C.I.: 19.4%, 38.2%). These false-positive error rates are much higher than those identified by [6] and [7]. However, they are roughly consistent with some of the CNM error rates reported in Table 1. For example, the 15.9% false-positive error rate for CNM_1 is not significantly different from the 13.1% error rate reported for U.S. examiners on the identical print pair in [19] ($Z = 0.54$, n.s.). However, the non-U.S. agencies in the present study offered fewer inconclusive opinions than did the examiners in [19] (14.4% vs 32.3%, $Z = -3.33$, $p < .001$).

The fingerprint agencies were encouraged to provide written comments following each conclusion reached on the test and most obliged. Table S1 presents the complete set of comments (translated to English by S.L.) provided by agencies that committed

a false-positive error on CNM_1. Table S2 provides a complete set of comments provided by agencies that committed a false-positive error on CNM_2. Some comments were edited lightly for consistency with English grammar rules.

Written comments were provided by 15 of the 17 agencies that committed false-positive errors on CNM_1 and by all 27 of the agencies that committed false-positive errors on CNM_2. These comments offer insight into why the false-positive errors may have occurred. Nine of the 15 commenting agencies (60.0%) that erred on CNM_1 referenced the large number of identical minutiae observed [8–15]. Similarly, 18 of the 27 agencies (66.7%) that erred on CNM_2 referenced the number of identical minutiae observed [9–15]. These results parallel those observed in [19]. It seems, then, that the presence of a large number of identical minutiae may play a key role in some examiners' decisions to call an identification. The agencies' comments also hint that many of those who committed false-positive errors may not have noticed important discrepancies between the latent and exemplar print. Only three of the 15 commenting agencies (20.0%) that erred on CNM_1 and eight of the 27 commenting agencies (29.6%) that erred on CNM_2 acknowledged finding any discrepancies within these print pairs. In a few instances, examiners reported similarities that did not exist. One agency reported a scar on both the latent and known print when no such scar was present. The agencies that did note discrepancies dismissed them as print distortion. However, there is little evidence that examiners can reliably differentiate between distortion and actual differences between prints [3,4,12,26]. Research is needed on this issue.

Finally, the written comments suggest that agencies that falsely identified CNMs were quite sure of their identifications (e.g., "They are from the same origin and cannot be reflected by another finger"; see Tables S1 and S2). These comments suggest that the great confidence examiners express on identifications in easy proficiency tests [27] may also exist on identifications (including false identifications) in very difficult tests.

6.1 | Discussion / Interpretation of results

The false-positive error rates of 15.9% on CNM_1 and 28.1% on CNM_2 are not consistent with the low error rates on non-CNM print pairs reported in [6] and [7]. The most likely explanation for this inconsistency is that the print pairs in the present study had more points of similarity and/or fewer points of dissimilarity than the print pairs used in [6] and [7]. It is also possible that the examiners in the present study were less skilled. However, this explanation is unlikely because our participants committed false-positive errors on CNM_1 at a rate similar to that observed by [19] on the identical CNM. Finally, we note that the PiAnoS software may have been new to many of the examiners and therefore may itself be an additional source of error. However, none of the agencies reported any frustrations with or difficulties using the software beyond the three agencies noted previously that were unable to use PiAnoS with the Internet Explorer browser.

If the high false-positive error rate in the present study is related to print difficulty [28,29], questions arise as to when such print pairs may be encountered in casework and what can be done to reduce the risk of error. Database simulations show that the number of CNMs increases faster than the chance that the suspect print will be in the database [30]. Verification of identifications by independent examiners will probably reduce the risk of error. Such verification is common in the United States, though there is some question about the independence because the verifying examiner typically knows the conclusion of the first examiner [26]. In some countries, independent verification is now required before database hits are classified as identifications.

7 | CONCLUSION

When a fingerprint examiner identifies a person as the source of a latent print, law enforcement personnel and legal decision-makers must know how much to trust that conclusion. References to an examiner's training, experience, and knowledge, and assurances that false identifications are "extremely rare" are inadequate. Science requires data that show the rate at which trained examiners reach accurate conclusions under conditions similar to those that exist in the case at issue [5]. Two well-designed studies provide error rate data that are relevant for estimating the accuracy of fingerprint analyses in many cases [6,7]. The tasks in these prior studies were not easy. Ulery et al. [6] deliberately selected non-mated exemplars in ways that increased the difficulty of the comparison process. Nevertheless, these studies do not indicate how high the false-positive error rate might be in more difficult CNM cases. The present study addresses this knowledge gap.

We obtained false-positive error rates of 15.9% and 28.1% on two fingerprint CNMs. This error rate is inconsistent with the popular notion that fingerprint evidence is nearly infallible. It is also more than two orders of magnitude larger than error rates reported in leading studies. This result suggests that, when the risk of CNMs is high, the probative value of a reported fingerprint match may be severely diminished due to an elevated risk of false identification.

There were features in the present study that are not common to previous fingerprint proficiency tests including mandatory participation, investigators from academia, and a focus on CNMs. The study also had limitations. First, the agencies and participating examiners knew they were being tested. Blind participation would be better to ensure that the analyses and reports provided are similar to those that are provided in actual casework [26]. Second, the CNMs were convenience samples, rather than random and representative selections of CNMs from database searches. We do not know if the CNMs used here were any more or less difficult than the CNMs examiners encounter in their casework. However, we note that a reviewer who has detailed knowledge about our two CNMs indicated that they are indeed "much more difficult" than most other CNMs. Therefore, we do not claim that the false-positive error rate observed here represents the error rate for fingerprint analyses when CNMs

are present or when databases are searched. A final reason to be cautious about the generality of the false-positive error rate that we identified is that different countries, agencies, and examiners may have different philosophies regarding the use of the "inconclusive" option. A greater reluctance to employ this option may increase the risk of both false-positive and false-negative errors. Nevertheless, the frequency of errors committed by the participating fingerprint agencies on this test hints that the error rate for CNMs may be substantial.

8 | FUTURE RESEARCH

We echo calls for research on forensic error rates [3,4,5,26] and suggest that research on fingerprint CNMs should be prioritized. There is reason to believe that CNMs will occur in large databases and that they pose a substantial risk for misidentification. Going forward, the present study should be replicated and expanded, if possible, using a representative selection of CNMs from database searches. Researchers should investigate conditions that can lead to CNMs (e.g., poor print quality, large database searches, prints from close relatives), as well as techniques that can be used to combat the confirmatory biases that prevent researchers from noticing and responding to CNM print dissimilarities [31,32]. The relationship between error and the availability and use of the "inconclusive" option is another area ripe for empirical investigation [33]. Research should also continue into how best to present error rates to forensic consumers. Studies suggest that people have trouble understanding and interpreting statistical evidence [34,35]. In particular, people often confuse reported false-positive error rates (i.e., the probability that an examiner would report an identification on a non-mated print pair) with the false discovery rate (i.e., the probability that an examiner's reported identification would be incorrect). Such transposition is prevalent in DNA cases and will likely arise in fingerprint cases when the testimony becomes more overtly quantitative. Finally, research on what jurors should be told about the risk of error must continue [26,36]. Finding the right balance is crucial. If jurors are told too little, they may assume that the risk of error is miniscule. If jurors are told that a CNM study reported a 21.7% false identification rate, they may assume that examiners err 21.7% of the time. Finding the right balance will help jurors and others assign the appropriate weight to fingerprint evidence.

REFERENCES

1. Garrett BL, Neufeld PJ. Invalid forensic science testimony and wrongful convictions. Va L Rev. 2009;95:1–97.
2. Koehler JJ. Intuitive error rate estimates for the forensic sciences. Jurimetrics. 2017;57:153–68. https://doi.org/10.2139/ssrn.2817443.
3. Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council. Strengthening forensic science in the United States: A path forward. Washington, DC: National Academies Press; 2009. https://doi.org/10.17226/12589.
4. Expert Working Group on Human Factors in Latent Print Analysis. Latent print examination and human factors: Improving the practice through a systems approach. Gaithersburg, MD: U.S. Department of Commerce, National Institute of Standards and Technology; 2012. doi: https://doi.org/10.6028/NIST.IR.7842.
5. The President's Council of Advisors on Science and Technology, Executive Office of the President. Report to the President - Forensic science in criminal courts: ensuring scientific validity of feature-comparison methods. 2016. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf. Accessed 1 Sep 2020.
6. Ulery BT, Hicklin RA, Buscaglia J, Roberts MA. Accuracy and reliability of forensic latent fingerprint decisions. Proc Natl Acad Sci USA. 2011;108(19):7733–8. https://doi.org/10.1073/pnas.1018707108.
7. Pacheco B, Cerchiai B, Stoiloff Y. Miami-Dade research study for the reliability of the ACE-V process: Accuracy & precision in latent fingerprint examinations (Document No 248534). Award number 2010-DN-BX-K268, 2014. https://www.ncjrs.gov/pdffiles1/nij/grants/248534.pdf. Accessed 1 Sep 2020.
8. Ausdemore M, Hendricks JH, Neumann C. Review of several false positive error rate estimates for latent fingerprint examination proposed based on the 2014 Miami Dade Police Department study. J Forensic Identif. 2019;69:59–81.
9. Organization of Scientific Area Committees, Friction Ridge Subcommittee. Response to call for additional references regarding: President's Council of Advisors on Science and Technology - Report to the President. 2016. https://www.nist.gov/fusion-search?s=psac-FR_PCAST. Accessed 1 Sep 2020.
10. Abraham J. Statistical models for the support of forensic fingerprint identifications [dissertation]. Sydney, Australia: University of Technology Sydney; 2017.
11. Langenburg GM. A critical analysis and study of the ACE–V process [dissertation]. Lausanne, Switzerland: Universite de Lausanne; 2012.
12. Office of the Inspector General, Department of Justice. A review of the FBI's handling of the Brandon Mayfield case. Washington, DC: Office of the Inspector General, Department of Justice; 2006.
13. Thompson GC. The potential for error in forensic DNA testing. GeneWatch. 2008;21:5–8.
14. Persinger M, Ericson L, Greene M. Latent fingerprint interoperability survey: A national study of automated fingerprint information systems (AFIS) maintained by law enforcement agencies. Washington, DC: U.S. Department of Justice, Office of Justice Programs; 2014.
15. Reaves B. Local police departments, 2007. Washington, DC: U.S. Department of Justice, Office of Justice Programs; 2011.
16. Dror E, Mnookin JL. The use of technology in human expert domains: Challenges and risks arising from the use of automated fingerprint identification systems in forensic science. Law Probab Risk. 2010;9:47–67. https://doi.org/10.1093/lpr/mgp031.
17. Collaborative Testing Services, Inc. Forensic testing program, latent prints examination. Report No. 9508. Sterling, VA: Collaborative Testing Services Inc; 1995.
18. Langenburg G, Champod C, Genessay T. Informing the judgments of fingerprint analyses using quality metric and statistical assessment tools. Forensic Sci Int. 2012;219:183–98. https://doi.org/10.1016/j.forsciint.2011.12.017.
19. Neumann C, Champod C, Yoo M, Genessay T, Langenburg G. Improving the understanding and the reliability of the concept of 'sufficiency' in friction ridge examination. Final Report. National Institute of Justice Award 2010-DN-BX-K267. Washington, DC: National Institute of Justice; 2013.

20. Liu S, Champod C, Wu J, Luo W. Study on accuracy of judgments by Chinese fingerprint examiners. J Forensic Sci Med. 2015;1(1):33–7. https://doi.org/10.4103/2349-5014.157908.

21. State v. Slater, 904 N.W.2d 144 (Wisc. App. 2017).

22. United States v. Baines, 573 F.3d 979, 990–91 (10th Cir. 2009).

23. Department of Justice, Office of Justice Programs. Fingerprint examiners found to have very low error rates. 2015. http://ojp.gov/newsroom/pressreleases/2015/ojppr02022015.pdf. Accessed 1 Sep 2020.

24. U.S. v. Herrera, 704 F.3d 480, 487 (7th Cir. 2013).

25. Clark JD. AVE-V: is it scientifically reliable and accurate? J Forensic Identif. 2002;52(4):401–8.

26. Thompson W, Black J, Jain A, Kadane J. Forensic science assessments: A quality and gap analysis – latent fingerprint examination. Washington, DC: American Association for the Advancement of Science; 2017. https://doi.org/10.1126/srhrl.aag2874.

27. Kelly S, Gardner BO, Murrie K, Pan DH, Kafadar K. How do latent print examiners perceive proficiency testing? An analysis of examiner perceptions, performance, and print quality. Sci Justice. 2020;60:120–7. https://doi.org/10.1016/j.scijus.2019.11.002.

28. Dror IE. Commentary on: the error in "error rate": Why error rates are so needed, yet so elusive. J Forensic Sci. 2020;65:1034–9. https://doi.org/10.1111/1556-4029.14435.

29. Kellman PJ, Mnookin JL, Erlikhman G, Garrigan P, Ghose T, Mettler E, et al. Forensic comparison and matching of fingerprints: Using quantitative image measures for estimating error rates through understanding and predicting difficulty. PLoS One. 2014;9(5):e94617. https://doi.org/10.1371/journal.pone.0094617.

30. Busey T, Silapiruti A, Vanderkolk J. The relation between sensitivity, similar non-matches and database size in fingerprint database searches. Law Probab Risk. 2014;13:151–68. https://doi.org/10.1093/lpr/mgu002.

31. Kassin SM, Dror IE, Kukucka J. The forensic confirmation bias: Problems, perspectives, and proposed solutions. J Appl Res Mem Cogn. 2013;2:42–52. https://doi.org/10.1016/j.jarmac.2013.01.001.

32. Dror IE. Biases in forensic experts. Science. 2018;360:243. https://doi.org/10.1126/science.aat8443.

33. Dror IE, Langenburg G. "Cannot decide": The fine line between appropriate inconclusive determinations versus unjustifiably deciding not to decide. J Forensic Sci. 2019;64:10–5. https://doi.org/10.1111/1556-4029.13854.

34. Koehler JJ. On conveying the probative value of DNA evidence: Frequencies, likelihood ratios and error rates. Univ Colo Law Rev. 1996;67:859–86.

35. Thompson WC, Kaasa SO, Peterson T. Do jurors give appropriate weight to forensic identification evidence? J Empir Leg Stud. 2013;10:359–97. https://doi.org/10.1111/jels.12013.

36. Thompson WC. How should forensic scientists present source conclusions? Seton Hall Law Rev. 2018;48:773–813.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.