Flawed forensics: Statistical failings of microscopic hair analysis

For 20 years, FBI forensic experts gave flawed testimony regarding microscopic hair analysis. **Jim Norton**, **William Anderson** and **George Divine** unpick their mistakes

In 1897, famed US Supreme Court justice, Oliver Wendell Holmes, Jr, said: "For the rational study of the law the black letter man may be the man of the present, but the man of the future is the man of statistics and the master of economics." Given his prescience on the growing importance of statistics, Holmes might not be surprised by the large number of legal cases where statistics play a key role. However, it is likely that he would be appalled by the misuse of statistics in many proceedings.

One such misuse involves the application of microscopic hair analysis in criminal investigations. In April 2015, the US Federal Bureau of Investigation (FBI) admitted that members

Image: SEInnovation/iStock/Thinkstock

of its forensic unit had given flawed testimony concerning microscopic hair analysis for over 20 years (see *Significance*, June 2015, page 3).

Hair is an obvious source of trace evidence in many crime scenes as it is frequently shed and easily transferred to clothes, sheets or carpets, or from one person to another. The first use of forensic hair analysis occurred in 1855 during the murder trial of John Browning. Hairs on a rope found in the defendant's home were visually compared to the victim's hair and were judged to be identical in colour and length.

Dr Sydney Smith, in 1934, was the first person to use microscopic hair comparisons in a murder trial by visually matching hair from the crime scene to the hair of the defendant. By the late 1970s, FBI laboratory personnel commonly used microscopic hair analysis as part of their toolkit in forensic investigations.

Limitations

There are several problems with using hair comparisons to connect a defendant with a crime scene, not least of which is determining the scientifically relevant characteristics of hair and the number of hair characteristics that must be compared to give reliable results. There is no law or standard addressing either of these aspects of hair comparison.

Microscopically, over 20 characteristics can be used to describe/identify a single hair, including pigment distribution, tip shape and shaft diameter. However, many of the characteristics are subjective, such as colour (is it yellow, yellow brown, or brown?), pigment distribution (uniform, peripheral, one side, clusters) and cortical texture (fine, medium, coarse). A re-examination of a hair, by the same or a different examiner, can result in different descriptions of hair characteristics.

Finally, and crucially, there are no population-based databases that contain subjects' hair characteristics, making it impossible to estimate the probability of any given hair characteristic.

In 2009, the National Academy of Sciences issued the report, *Strengthening Forensic Science in the United States: A Path Forward*.[1] The report discussed the use of forensic science by law enforcement, devoting seven pages to hair comparison. It concluded, based on the problems identified above, that microscopic hair comparisons simply "cannot uniquely identify one person".

That same year, Cary Oien, chief of the FBI's Trace Evidence Unit, described the proper way to write up and interpret the results of a hair comparison.[2] According to Oien, only three general conclusions can be reached from an examination: exclusion, no conclusion, or association. When the differences between hair from the crime scene and hair from the subject are dramatic, the subject can be excluded from the set of people who could have left the hair. When the suspect's hair exhibits some similarities to the donor's hair, but also exhibits some slight microscopic differences, then no conclusion can be drawn. If the suspect's hair exhibits the same microscopic characteristics as the donor's hair (from the crime scene), there is an association and one concludes that the hair is consistent with the same source as the known sample.

H. James (Jim) Norton is professor emeritus of biostatistics, Carolinas HealthCare System, Charlotte, NC. He has been an expert witness or consultant in eight legal cases, including six that involved forensic evidence

William E. Anderson, MS, JD is a statistician at Dickson Advanced Analytics, Carolinas HealthCare System, Charlotte, NC. He also practised law for over 10 years, specialising in cases involving medical and scientific issues

George Divine is a senior research biostatistician at Henry Ford Hospital, Detroit, Michigan

Oien cautions that we do not know the proportion of people in the general population who have these same hair characteristics, or the probability of a coincidental match of two hairs. Nevertheless, laboratory technicians and expert witnesses often give probabilities or pseudo-probabilities when testifying in criminal cases.

Evidence in court

In a 2012 article in the *Washington Post* (wapo.st/21dnZdG), reporter Spencer Hsu described the flawed hair analysis that led to the conviction of three men who were later exonerated by DNA testing. One of the men, Kirk Odom, served 22 years in prison for murder.

In another well-publicised case, Jimmy Ray Bromgard was exonerated in 2002 of a rape he did not commit. In that case, an eight-year-old girl was raped in 1987. Based on her recollection, a sketch was made of the perpetrator, and a police officer thought it resembled Bromgard.

Bromgard agreed to take part in a line-up, which was filmed. According to the website of the Innocence Project (bit.ly/21do5SG), an organisation that works to overturn wrongful convictions: "In the live proceedings, the victim picked out Bromgard but was not sure if he was the right man. After the victim was shown the videotaped footage of Bromgard, she said she was '60%, 65% sure'. When asked at trial to rate her confidence in the identification without percentages, she replied, 'I am not too sure'."

The key evidence against Bromgard was hair left at the scene of the crime that was compared to Bromgard's hair. The prosecution's forensic expert testified that the probability the hair came from someone else was one in 10 000. Bromgard was convicted and spent 14 years in jail, until DNA testing showed that he could not have been the person who assaulted the young girl.

Prompted by such reports of injustice, a consortium was formed in 2013 to review approximately 3000 criminal cases in which the FBI used microscopic hair analysis to help convict a defendant. The group included members of the Innocence Project, the National Association for Criminal Defense Lawyers, the law firm Winston & Strawn LLP, the FBI and the Department of Justice.

As of April 2015, the consortium had reviewed 269 transcripts involving microscopic hair analysis and had

Microscopic hair comparisons cannot uniquely identify one person. We do not know the proportion of people in the general population who have the same hair characteristics

determined that 96% of the transcripts had at least one testimonial error. In the 35 capital cases with suspect testimony, nine defendants were executed, and five more died while on death row. This review led to the FBI's admission of flawed testimony in numerous criminal cases from the 1980s until 1996, when it began using mitochondrial DNA (mtDNA) analysis for hair comparisons.

The problem with probabilities

Both fact witnesses and expert witnesses have misused probabilities when interpreting hair comparison results in legal cases. Laboratory personnel, most of whom are not formally trained in probability theory, have been known to invent probabilities when testifying on hair comparison results.[3] For example, a North Carolina lab technician testified: "If you pick an individual at random off the street, there is one out of a thousand chance that the unknown hair would match or would be consistent with that person's hair." There is no scientific basis for such a claim. Similar baseless statements made at trial include:

- "Reddish-yellow hair [is found in] about 5% of the population."
- "I have looked at thousands of hair standards over the course of my work and I haven't seen any that have the same range of physical characteristics yet." (This is an example of a pseudo-probability. The jury may interpret this statement as "the probability the hair comes from someone other than the defendant is one in thousands".)
- "There is one chance, perhaps for all we know, in 10 million that it could [be] someone else's hair." (Prosecutor's summary to a jury)

In the Bromgard case, the prosecutor's forensic expert testified that he found a match between the defendant's scalp and pubic hair and hair from the crime scene. Without offering a basis, he stated that the probability of a match for scalp hair was one in a hundred, and gave the same probability for a pubic hair match. He opined that, since the hairs were from different parts of the body, they were independent. Then, using the multiplication rule for independent events, he concluded that the chance that the hair came from someone other than the defendant was one in 10 000.

As in other cases where the multiplication rule was misused, the expert witness multiplied probabilities (to get a much smaller probability) without presenting any evidence that the events are independent. Indeed, a person's scalp hair and pubic hair have similarities (such as colour), so that a match on a person's scalp hair and a match on their pubic hair could not be considered independent events.

In another case, an expert witness testified that it would not be unusual to have to examine 4500 strands of scalp hair to get a match with any particular hair. The basis offered for the probability of 1/4500 was a study by Gaudette and Keeping.[4] Others have criticised this study as flawed[5], however, and have cautioned that its results should never be used in a legal case.

Indeed, hair may not be as unique to an individual as proponents of microscopic hair analysis believe. In Gaudette and Keeping's study, 13 of 100 individuals had a strand of scalp hair that was indistinguishable from that of another individual. For pubic hair, the number was 25 of 60 individuals. In 2002, an FBI agent performed an mtDNA analysis on 80 pairs of hair that his laboratory had previously been asked to evaluate microscopically. Despite the fact that all 80 pairs had been judged as being associated (a match), mtDNA proved that in 9 pairs (11%), the hairs were not from the same individual.

Where are the experts?

Federal courts and many state courts use the so-called *Daubert* factors[6] or similar criteria to determine whether a scientific theory or technique is sufficiently reliable to be admitted in court.[7] The five factors are: whether the theory or technique can be and has been tested; whether it has been subjected to peer review and publication; the known or potential error rate; whether it is generally accepted in the scientific community; and whether standards and controls regarding the theory or technique exist and are maintained.

Laboratory personnel, most of whom are not formally trained in probability theory, have been known to invent probabilities when testifying on hair comparison results

Notwithstanding the *Daubert* requirements, courts generally have found microscopic hair analysis reliable. An exception was the case of Ron Williamson, who had been convicted of murder in 1988 and sentenced to death. After several appeals, an Oklahoma district court in 1995 granted Williamson a new trial.[8] In its opinion, the court noted that, under "the guidelines of *Daubert*, this court has found an apparent scarcity of scientific studies regarding the reliability of hair comparison testing. The few available studies reviewed by this court tend to point to the method's unreliability." DNA evidence later demonstrated that the prosecution's main witness, who testified that he saw Williamson at the scene of the crime just prior to the time of the murder, was in fact the murderer. Williamson was exonerated and released from prison in 1999.

Similarly, the Court of Appeals of North Carolina ruled that a hair expert's testimony in *State* v. *Faircloth* was in error.[9] The court found that he overstated the reliability of microscopic hair comparisons when he told the jury that it was "impossible" for someone other than the defendant to have been in contact with the crime area and the victim's person. Short of using the

word "impossible", however, exaggerated claims regarding hair comparison testing have been admitted routinely in court.

Despite prosecutorial misstatements regarding microscopic hair analysis, criminal defendants have underutilised expert witnesses to rebut such testimony. In a review of 137 persons who were convicted of serious crimes in which DNA analysis led to post-conviction exoneration, a defence expert testified in only 19 cases.[3] As with judges and juries, many defence attorneys do not understand the problems underlying microscopic hair analysis.

Even when the defence lawyer recognises the need for an expert witness, however, the defence may not be allowed to hire one. For example, when a public defender's office represents an indigent person, the judge typically determines whether the defendant is entitled to an expert witness. If the judge believes the evidence against the defendant is overwhelming, or if he is concerned about limited financial resources, he may disallow the use of an expert witness.

Summing up

Microscopic hair analysis is a flawed forensic technique, its deficiencies exacerbated when coupled with dubious statistical conclusions proffered into testimony. A sample *Daubert* motion drafted by the Innocence Project of New Orleans (and posted online at bit.ly/21dr7q0) reflects this view. In short, it makes clear that microscopic hair analysis does not pass muster under *Daubert* and has no place in a courtroom.

In contrast to microscopic hair comparison, DNA testing is the gold standard for identifying individuals based on trace evidence left at a crime scene. Unlike in microscopic hair analysis, a person's DNA essentially is fixed and unique, and DNA testing is replicable, laboratory verifiable, and based on large databases. Perhaps it is no surprise, then, that since 1989, 74 people who were convicted of serious crimes, in large part due to microscopic hair comparisons, were later exonerated by post-conviction DNA analysis (see box). ■

References

1. National Research Council, National Academy of Sciences (2009) *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: National Academies Press. 1.usa.gov/1CXBCpr

2. Oien, C. T. (2009) Forensic hair comparison: Background information for interpretation. *Forensic Science Communications*, **11**(2).

3. Garrett, B. L. and Neufeld, P. J. (2009) Invalid forensic science testimony and wrongful convictions. *Virginia Law Review*, **95**(1), 1–97.

4. Gaudette, B. D. and Keeping, E. S. (1974) Attempt at determining probabilities in human scalp hair comparison. *Journal of Forensic Sciences*, **19**(3), 599–606.

5. Aitken, C. G. G. (1995) *Statistics and the Evaluation of Evidence for Forensic Scientists*. Chichester: John Wiley & Sons.

6. *Daubert* v. *Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1993)

7. Giannelli, P. C. (2011) Daubert and forensic science: The pitfalls of law enforcement control of scientific research. *University of Illinois Law Review*, **2011**(1), 53–90.

8. *Williamson* v. *Reynolds*, 904 F. Supp. 1529 (E.D. Okla. 1995).

9. *State* v. *Faircloth*, 394 S.E.2d 198 (1990).

The exonerated

The Innocence Project report, *Not a Strand of Evidence* (bit.ly/1QbiqJf), documents the cases of 74 people who were convicted using microscopic hair comparisons and who were later exonerated by DNA testing.

Of these cases:

- 20 states and the District of Columbia are represented
- 45% are from Illinois, New York or Oklahoma
- 73 are men, one woman
- Total years served in prison: 1056 (mean 14.3 years, range 3–36.5 years)
- In 41% of cases, the real perpetrator has been found.

Primary crime:

- Murder or homicide – 62%
- Rape or sexual assault – 37%
- Robbery – 1%.

Factors contributing to conviction, other than unvalidated/improper forensic science:

- Eyewitness misidentification – 64%
- False confessions or admissions – 36%
- Jail house informants and 'incentivised' witnesses – 24%
- Inadequate defence – 8%
- Government misconduct – 1%.

21597185/iStock/Thinkstock