

# An Objective and Statistical Approach to Microscopic Human Hair Comparison: A Laboratory Exercise for the Forensic Science Undergraduate and Graduate Student

Emma Redman<sup>1</sup>, Casey Rech<sup>1</sup>, Isabel Sandone<sup>1</sup>, Victoria Echternach<sup>1</sup>, Lawrence Quarino<sup>1\*</sup>

<sup>1</sup>Department of Chemical, Physical, and Forensic Sciences, Cedar Crest College, 100 College Drive, Allentown, PA 18104

\*corresponding author: laquarin@cedarcrest.edu

**Abstract:** The following introduces a new approach to teaching microscopic hair examination in an academic instructional laboratory for forensic science undergraduate and graduate students. In the exercise, students are asked to determine the likelihood ratio of test hairs to assess the probability of encountering a hair with similar characteristics. Instead of relying on qualitative subjective assessment of morphological characteristics, students use two quantitative and objective parameters, namely diameter and color to characterize test hairs. With the use of software measurement tools, the diameter of each hair was measured in 3 locations along the hair shaft toward the middle of the hair and five RGB (red/green/blue) values were recorded at different points in the cortex approximately 3 um from the edge of the hair. Values are compared to a constructed hair database created from collected hairs vacuumed from heavily trafficked areas such as dining halls and lecture halls to determine a random match probability. A 95% upper bound confidence interval was determined from each random match probability and the reciprocal of this value was used to calculate a likelihood ratio which ranged from approximately 100 to 400 for randomly collected hairs. It is hoped that an important learning outcome of this exercise is that forensic science students will develop an awareness of the importance of providing statistical meaning to forensic science inclusions thus reducing the potential for scientific information to be misconstrued. This approach differs from most academic laboratory exercises of this nature which focus exclusively on matching unknowns to a closed set of standards.

**Keywords:** microscopic hair comparison, likelihood ratios, RGB color format, diameter

## Introduction

Many full service forensic science laboratories have scaled back trace evidence services. Reasons for this include slower analysis time leading to longer throughout of cases, a lack of requests for trace evidence examinations, a lack of individualism potential, and difficulty in hiring personnel with expertise in trace evidence examinations. In the age of forensic DNA analysis where a biological sample essentially can be linked to an individual with near certainty, it is not surprising that many district attorney offices devalue results of many trace evidence examinations which typically does not determine the unique source of the evidence. The situation with forensic microscopic hair examination is even direr considering that the reliability of such examinations is often called into question by the legal and scientific community. Although hair examination has been accepted in US courts for decades, it has been described by legal scholars as “snake oil” (1) and many cases have been reported where forensic hair examination is alleged to have contributed to false convictions (2,3,4). Criticisms such as these led the Federal Bureau of Investigation (FBI) in conjunction with

the United States Department of Justice and National Association of Criminal Defense Lawyers to undertake a systematic review of past FBI laboratory casework involving forensic hair examination. The results of this investigation were staggering. In the years prior to 2000, the study revealed that FBI trace evidence scientists routinely provided erroneous statements regarding hair examinations in laboratory reports and in testimony (5). Concerns about the reliability of microscopic hair examination have caused many forensic laboratories to remove forensic microscopic hair examination from their trace evidence services. The questions of reliability stem largely from the subjective nature of microscopic hair examination. The subjective determination whether two hairs could have originated from the same source involves the comparison of many phenotypic and morphological characteristics such as the medulla, cuticle, cortical fusi, and pigment granules. Subjective analysis should not be synonymous with unreliability and many studies have demonstrated the reliability of microscopic hair comparison. Strauss, for instance, reported no false inclusions or exclusions occurred in 4,900 comparisons (6). This study seemed to confirm the earlier work of

Gaudette and Keeping which found that out of 366,630 pairwise comparisons only nine pairs of hair were indistinguishable (7). A more recent study provided similar comparative date (8). In 2002, Houck and Budowle found that mitochondrial DNA only excluded 9% of positive microscopic hair comparisons (9). Although it appears that the forensic community has reached a consensus that microscopic hair examination cannot be used to uniquely identify an individual, it is still nonetheless reasonable to conclude that accurate and reliable comparative analysis of hair morphology is possible (if for no other reason than for exclusion purposes or identifying possible hair matches that may be resolved by DNA) but requires years of experience to achieve the level of expertise required.

How then can an academic program provide laboratory instruction to college and graduate forensic science students in this type of comparison that will emphasize a scientific approach and not be based on training and experience? The answer may lie in limiting hair characteristics to those that can be measured quantitatively and not assessing on the ability to match an unknown to a closed set of standards which by current practice might be pointless. With quantitative data, the possibility of understanding the meaning and significance that two hairs are microscopically indistinguishable exists if there is a larger population of hair to compare it to. If hair similarity can be assessed statistically, it likely removes the notion of the determination of a unique origin.

This laboratory exercise attempts to do this using two parameters that can provide quantitative data, namely diameter and color which can be recorded digitally with most imaging software programs. This exercise, which currently is being offered as part of an undergraduate trace evidence course in a Bachelor of Science in forensic science program, utilizes the RGB (red/green/blue) color format which has been shown to be helpful in differentiating brown Caucasian hairs from different individuals (10).

The meaning of hair similarity between hairs taken from different people can then be assessed through comparison of test hairs to a database of hairs having diameter and color measurement values. Having determined the frequency of a diameter and color combination (random match probability), an upper bound 95% confidence interval can be generated which then can be converted into a likelihood ratio assessing the rarity of the hair characteristics tested. This is similar to the approach applied in reporting population frequency of DNA haplotype matches.

## Methods

### *Development of Hair Database*

A trace evidence vacuum (**FIGURE 1**) was used to vacuum common areas around campus such as dining halls, lounges and lecture halls. Collection canisters were emptied and presumed hairs were removed and mounted onto microscope slides with large cover slips with DPX mounting media ( $n_D - 1.521$ ; Sigma Aldrich Prod. No.44581). All non-human hairs and fibers were discarded (based on microscopic characteristics and morphology), and each human hair was examined at 200X using an Olympus BX53 polarizing light microscope with CellSens® Image Capture Software (Olympus, Center Valley PA) under Kohler illumination and standardized lighting conditions (**FIGURE 2**).

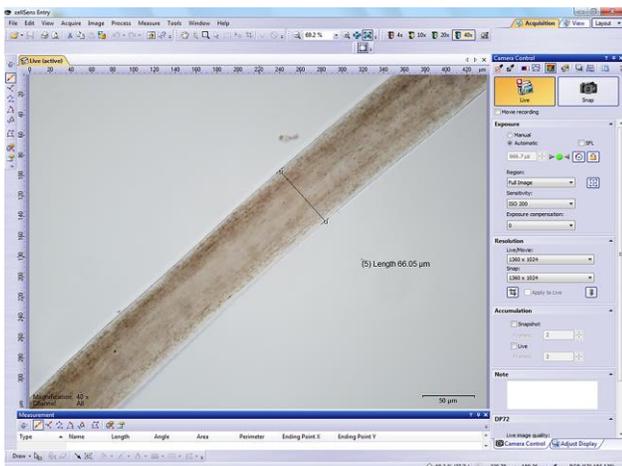
An image was captured of the middle portion of each of the hairs and the line measurement tool on the software was used to take the diameter across the hair in five locations. The software also allows for color to be measured quantitatively using the red-green-blue (RGB) color system which provides numerical color values for each color. To account for the variability and uncertainty of RGB values across the cortex of a human hair, measurements were taken approximately 3  $\mu\text{m}$  from the edge of the hair at 5 locations in the middle of the shaft of the hair (**FIGURE 3**).



**FIGURE 1:** Top is a trace evidence vacuum; bottom is collection filter showing collected material on filter.

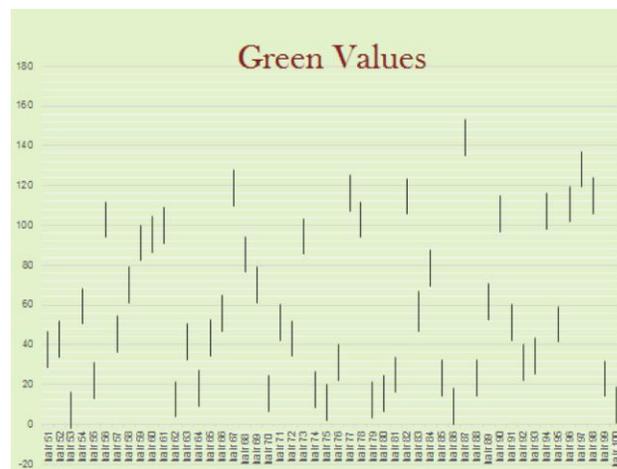


**FIGURE 2:** Olympus BX53 polarizing light microscope with computer monitor.



**FIGURE 3:** Image capture of hair magnified at 200x showing measurement tool across the diameter of hair. RGB values are recorded at the point of the mouse cursor (not shown). RGB values are shown in lower right hand corner of screen; numerical values for red, green, and blue are recorded respectively.

For each hair, mean and standard deviation values were generated for diameter and each color value. The mean standard deviation value of all the hairs combined in the database (N=250) was calculated and used to create a ± “bin” around the mean of each parameter for every hair. **FIGURE 4** displays bins for 50 selected hairs for green color value. The generation of bins made comparisons of test hairs to the database possible. The data base was compiled in Microsoft Excel® with no special software used.



**FIGURE 4:** Part of green color database showing bins for individual hairs. Green color values are listed on Y axis, hair identification number is the X axis.

*Test Hair Comparison to Database and Statistical Analysis*

After Institutional Review Board approval, four test hairs (blond, light brown, dark brown, and red in color respectively) were collected from subjects and five replicate measurements of diameter and RGB values were recorded. The mean value for each parameter was generated and compared to the database to determine the frequency of the diameter and each color value of each test hair (essentially if the mean value for any parameter fell within a bin it was considered “similar”). A random match probability for each test hair was generated by multiplying the frequency of occurrence of the diameter and the three color values in the database (pairwise correlation analysis of each of the color databases was previously performed showing no relationship between colors). An upper bound 95% confidence interval was generated from the random match probability using:

$$\text{EQUATION 1: } p+1.96[p(1-p)/N]^{1/2}$$

where p is the random match probability and N is the number of hairs in the database. The reciprocal of the upper bound confidence interval value was taken as the likelihood ratio. All calculations were performed manually.

Likelihood ratios were used in this exercise because they (and by extension Bayesian statistics) are commonplace in forensic DNA profiling and their use has been suggested for many types of forensic evidence including trace evidence (11). The generation of a likelihood ratio involves the ratio of probabilities of competing hypotheses. In this exercise, hypothesis #1 (numerator) is considered the prosecutor’s hypothesis and is given the value of 1 because the prosecutor is believed to be offering the evidence as proof that the hair is from a particular source to the exclusion of all others. Conversely, hypothesis #2 (denominator) is considered the defense

attorney's hypothesis which states that the hair came from some other source than the alleged source. In this exercise, the probability of the defense attorney's hypothesis is the frequency of the hair characteristics in the generated database (denoted as the random match probability).

## Results

When compared to the database, the frequency of occurrence of the mean diameter of the four test hairs ranged from 0.100-0.188 mm. The frequency of mean color values of the four test hairs ranged from 0.152-0.232 (red), 0.132-0.308 (green), and 0.112-0.352 (blue) respectively. Multiplying the frequency of the diameter by the frequency of each color value produced random match probabilities in the  $10^{-3}$  to  $10^{-4}$  range. With the upper bound confidence intervals, all probabilities were in the  $10^{-3}$  range. Subsequent calculation of the likelihood ratio for each hair produced the following results: 336 for the blond hair, 216 for the lighter of the brown hairs, 164 for red hair and 106 for the darker of the brown hairs.

## Discussion and Conclusion

The fundamental educational benefit of this exercise is that forensic science students will develop an awareness of the importance of providing meaning to forensic science inclusions. Even for those who believe that microscopic hair comparison can never justify statements about identity, the often used phrase "consistent with," "could be the source of," or "cannot be excluded" is also problematic. The now defunct National Commission on Forensic Science recognized the danger of such language because it allows jurors who hear such testimony to simply make their own interpretation on what "consistent with" means (12). It is not unreasonable to believe that without proper context, a strong possibility exists that at least some jurors will believe that the proposed source of the evidence is more likely than not.

Some believe that the complexity involved in forensic hair comparison make the possibility of developing a statistical framework likely impossible (13). It is not analogous to forensic DNA profiling because science has an understanding of how Mendelian genetics works. Genetic variability is understood and can be quantitated. Conversely, at least to this point, the extent of variability of human head hair is not known and perhaps is unknowable. The exercise presented does not suggest that this approach is transferrable to everyday forensic science practice. Given the extent of the variability of hair, 250 head hairs (assumed head hairs) in a database taken from one geographical area may not be enough for extrapolation to a larger population.

Even still, given the likelihood ratios obtained from the test hairs will demonstrate to students that false inclusions are possible. Being able to somehow quantitate

that likelihood is vital for a deeper understanding of the meaning of evidence particularly within the framework of testing conducted. This exercise only used two variables which were measurable. If more parameters were used, the statistical likelihood of an inclusion would surely be higher. Students also need to be made aware that evidence in investigations is often cumulative and a statistical likelihood in one case may have different meaning than in another case.

The exercise could be expanded to include a comparison of test hairs to a set of hair exemplars using the methodology described. Mean and standard deviation could be calculated for test and exemplar hairs (data from exemplar hairs from each set can be grouped together) and comparisons made between test hairs and sets of exemplar hair based on the data. An inclusion between test hair and set of exemplars occurs when  $\pm$  standard deviation values around a mean overlap between test hair and an exemplar(s) set at all four parameters. At this point, the likelihood ratio (LR) of the inclusion could be determined from the database and students should be instructed to provide the following conclusion:

*Given the available information, the probability of these hair comparison results is **LR** times greater if the prosecution's proposition is true than if the defendant's proposition is true.*

The authors piloted this exercise with exemplars from five individuals all with the same shade of blond hair. The correct outcome was achieved.

In order to perform this exercise, a microscope with imaging software with RGB (or other color format) capability and measuring tools needs to be available. Also, a database of hairs with RGB and diameter values needs to be created ideally in a program that is searchable. Once established, however, the database can be used repeatedly.

## Acknowledgements

This work was performed as part of an incoming first-year student science research program at Cedar Crest College called Aspire. The authors in particular wish to thank Dr. Elizabeth Meade, President of Cedar Crest College, and Brianna Gregory who served as a student mentor to the participants.

## References

1. Smith CAS, Goodman PD. Forensic hair comparison analysis: nineteenth century science or twentieth century snake oil? *Col Hum Rts LR* 1996;27:227.
2. LaPorte GM. Wrongful convictions and DNA exonerations: understanding the role of forensic science. *National Institute of Justice Journal* 2017.

<https://nij.ojp.gov/topics/articles/wrongful-convictions-and-dna-exonerations-understanding-role-forensic-science>.  
Accessed: December 30, 2021

3. Forensic failures: three men, three hairs, three wrongful convictions. Locard's Lab.  
<https://locardslab.com/2016/03/03/forensic-fails-three-men-three-hairs-three-wrongful-convictions/>.  
Accessed: December 30, 2021
4. Giannelli PC. Wrongful convictions and forensic science: the need to regulate crime labs. *NCLR* 2007;86(1):163-285.
5. FBI testimony on microscopic hair analysis contained errors in at least 90 percent of cases in ongoing review. *FBI News* 2015.  
<https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review>. Accessed: December 30, 2021
6. Strauss MAT. Forensic characterization of human hair. *The Microscope* 1983;31:15-29.
7. Gaudette BD, Keeping ED. An attempt at determining probabilities in human scalp hair comparison. *J Forensic Sci* 1974;19:599-606.
8. Wickenheiser RA, Hepworth DG. Further evaluation of probabilities in human scalp hair comparisons. *J Forensic Sci* 1990;35:1323-1329.
9. Houck MM, Budowle B. Correlation of microscopic and mitochondrial DNA analysis of hairs. *J Forensic Sci* 2002;45:964-967.
10. Mills M, Bonetti J, Brettell TA, Quarino L. Differentiation of human hair by colour and diameter using light microscopy, digital imaging and statistical analysis. *J Microsc* 2017;270:27-40.
11. Buzzini P, Curran JM. Interpreting trace evidence. In: *Handbook of Trace Evidence Analysis*. Desiderio VJ, Taylor CE, Nic Daeid N (eds), Wiley and Sons, 2020, pp. 421-454.
12. National Commission of Forensic Science. Inconsistent language. National Institute of Standards and Technology 2015.  
<https://www.justice.gov/archives/ncfs/file/47784/>.  
Accessed: December 30, 2021
13. Watkins T, Bisbing RE, Houck M, Betty B. The science of forensic hair comparisons and the

admissibility of hair comparison evidence: Frye and Daubert considered. *Modern Microscopy*. The McCrone Group.

<https://www.mccrone.com/mm/the-science-of-forensic-hair-comparisons-and-the-admissibility-of-hair-comparison-evidence-frye-and-daubert-considered/>.

Accessed: December 30, 2021