Likelihood Ratio Model in AFIS

Testing a New Tool to Measure the Weight of Latent Evidence

Eric Ray, CLPE
Technical Project Leader, IDEMIA Identity and Security North America

INTRODUCTION

Eric Ray, CLPE

- **Background**
  - 12 years as a Latent Print Examiner for AZ and I3
  - B.S. in Biochemistry and Biology from University of Arizona
  - Member of OSAC Friction Ridge Subcommittee
  - Editorial Board of the Journal for Forensic Identification
  - Training to Latent Print Examiners across the US, Europe, and Asia
  - Co-host of the Double Loop Podcast (250+ episodes over 9+ years)
- **IDEMIA**
  - Technical Project Leader – Product Management
  - ABIS products – MBIS and STORM
  - Design of new web-based ABIS
  - Convey user needs and recommend product improvements
  - Training and troubleshooting

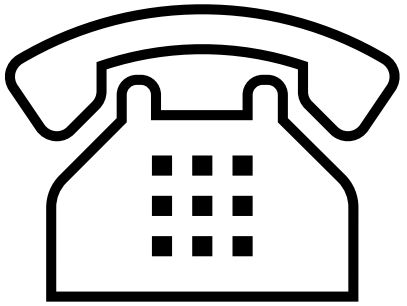
AGENDA

- Introduction
- Statistical Basics
- Theoretical Model
- Practical Model
- Test Results
- Questions

How Do We Arrive at an Identification?

- “Friction ridge identification is established through the **agreement** of friction ridge formations, in sequence, having sufficient uniqueness **[discriminating power]**”
  - *Quantitative-Qualitative Friction Ridge Analysis* by David Ashbaugh – 1999
- In order to reach an ID:
  - The two impression must have sufficient correspondence
  - The latent print must have sufficient discriminating power
- For every examination, we mentally measure both:
  - How similar are the features in these two impressions?
  - How specific or generic is this set of features?

Calls for a Statistical Approach



- **National Research Council** – *Strengthening Forensic Science in the United States*
  - More research is needed regarding the discriminating value of the various ridge formations and clusters of ridge formations. It also would lead to a good framework for future statistical models and provide the courts with additional information
- **NIST** – *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach*
  - [The federal government should] facilitate the validation of probabilistic models and other statistical research
  - The latent print examiner community should expand the training of examiners… to properly utilize the output of probabilistic models.
- **PCAST** – *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*
  - [E]xpert should report the probative value of the observed match based on the specific features observed in the case.
  - An expert should not make claims or implications that go beyond the empirical evidence and the applications of valid statistical principles to that evidence.
- **OSAC – Friction Ridge Subcommittee** – *Needs Assessment*
  - Expanded research on different statistical models that can be used in association with a friction ridge comparison… for quantification of evidence

Benefits of a Statistical Approach

- **Inform and support the opinions of Latent Print Examiners**
  - Will not replace ACE-V or the expert opinions of LPEs
- **More resilient to admission challenges**
  - NC v. McPhaul – LPE must document and testify to the specific features in this examination
  - MD v. Abruquah – FTE should not be permitted to offer an unqualified opinion on bullet comparison
- **More objective measurement of evidence**
  - Not solely based on examiner opinion (“Just trust me!”)
- **Quantifies the weight of the evidence**
  - Specific to the evidence *in this case*
  - Differentiates “barely ID” vs. “overwhelming ID”
  - Assists in estimating the risk of error for this comparison
- **Focuses additional Quality Assurance measures**
  - Highlight challenging or risky examinations for blind verification or further review

Statistical Basics

- Likelihood ratio (LR) (or a Bayes factor) can calculate the weight of the evidence
- LR starts with stating two, mutually exclusive propositions
  - H_1 – Prosecution proposition – Latent print originated from Mr. X
  - H_2 – Defense proposition – Latent print did not originate from Mr. X, but from a different person
- What is the probability of observing the evidence given the prosecution proposition?
  - If the correspondence is very strong (almost an overlay), then this probability increases.
  - If the correspondence is weak and requires “making it fit”, then this probability decreases.
- What is the probability of observing the evidence given the defense proposition?
  - If the corresponding features are very generic, then this probability increases. (Maybe it’s a CNM?)
  - If the corresponding features are very specific, then this probability decreases. (Not likely a CNM)
- This lines up very well with what we do during comparison
  - How similar are the features in these two impressions?
  - How specific or generic is this set of features?

Statistical Basics

$$Likelihood\ Ratio = \frac{Similarity}{Rarity}$$

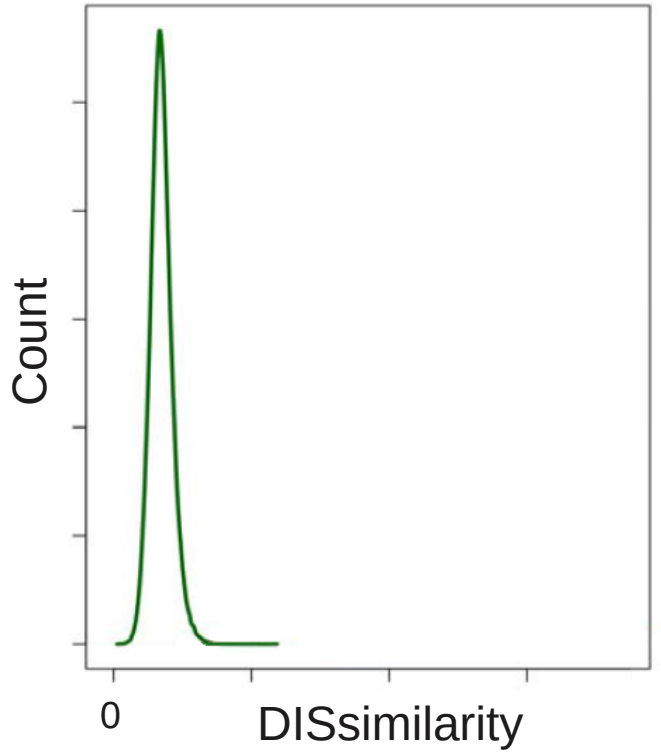
- Measuring similarity (or dissimilarity)
  - Similarity – an exact overlay of two impressions produces the highest value
  - Dissimilarity – an exact overlay of two impressions = 0 dissimilarity; higher values = more differences
  - High similarity or low dissimilarity corresponds to a high probability for the prosecutor proposition
- Measuring rarity or specificity
  - Search a large database representative of the relevant population
  - If there is high similarity (or low dissimilarity) to non-matching prints, then features are not very specific
  - If very few have high similarity (or low dissimilarity) to the latent print, then features are relatively specific

What Are We Looking For?

- **Something that answers our two questions**
  - How similar are the features in these two impressions?
  - How specific or generic is this set of features?
- **Needs to be based on the latent print *in this case***
  - Not based on averages or samples
- **Ratios need to use the same terms**
  - The Prosecution and Defense propositions must use the same process so we can create a valid Likelihood Ratio
- **Tested on Close Non-Matches**
  - Not based on obvious non-matches
- **Available**
  - Law enforcement must be able to purchase and deploy

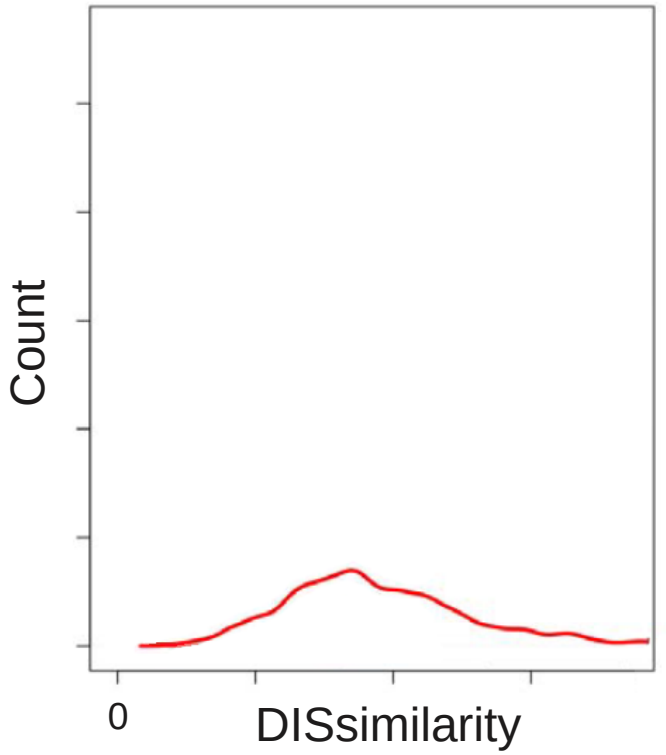
Theoretical Model

- We've reached an Identification conclusion, and want to test this hypothesis
  - We ask Mr. X to make a million latent prints using the same area of skin as the Latent Print
  - Measure and graph the DISsimilarity of the Latent Print to each of the million latent prints
- If the Known Print truly is the source of the Latent Print, then most of the samples will have low dissimilarity
  - Samples with more distortion will have higher dissimilarity
  - Samples that almost perfectly overlay with the Latent Print will have dissimilarity close to 0
- If we obtain even more sample latent prints from Mr. X, then the part of this curve that approaches zero dissimilarity estimates the Prosecution Proposition



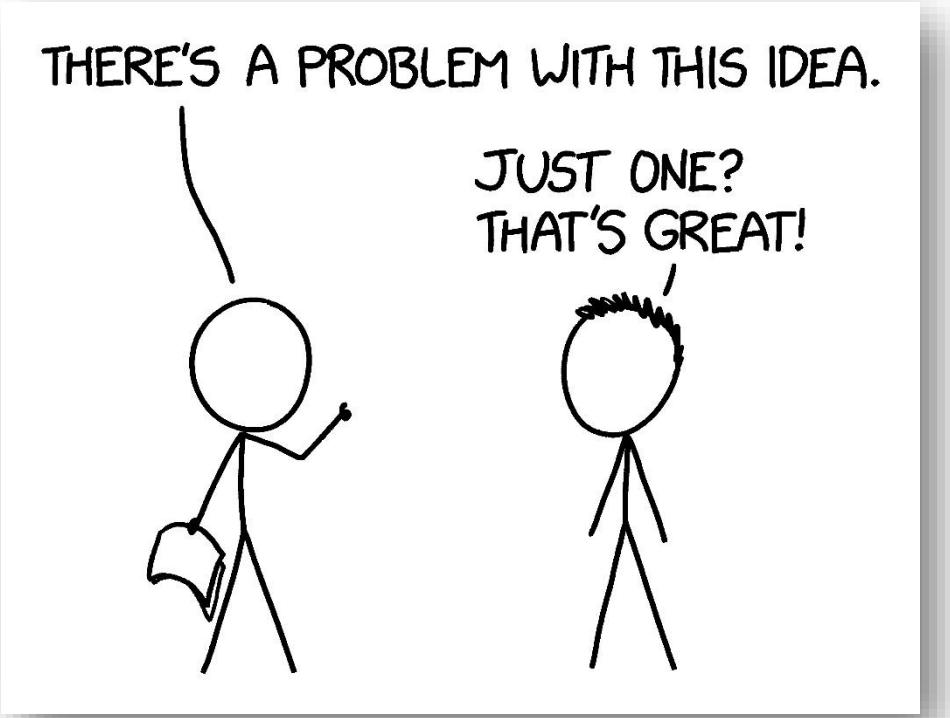
Theoretical Model

- Now let's consider the Relevant Population
  - We next ask everyone that could have ever left the latent print to each make a million latent prints using the most similar area of skin possible
  - Measure and graph the DISsimilarity of the Latent Print to each of the millions and millions of latent prints
- Most of the new latent prints will be very dissimilar to the Latent Print
  - Some Close Non-Matches will have low dissimilarity
- If the Latent Print is very specific with rare features, then this distribution will be shifted to the right
- If the Latent Print is very generic and has low specificity, then this distribution will be shifted to the left
- If we obtain even more sample latent prints from the Relevant Population, then the part of this curve that approaches zero dissimilarity estimates the Defense Proposition



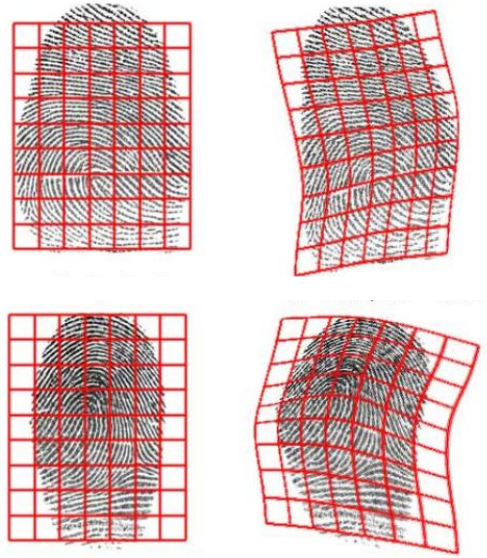
Theoretical Model

- **Is this what we're looking for?**
  - Measures the (dis)similarity between Latent Print and Known Print
  - Measures the specificity or rarity of the Latent Print features
  - Both distributions are based on the Latent Print *in this case*
  - Same process used for both propositions so we can create a valid ratio
- **Problems**
  - It's unlikely that we can get the suspect's cooperation to make millions of latent prints
  - And the rest of population of the world probably has better things to do
  - And it might take awhile for us to do all those comparisons
- **How can we make this work?**



Practical Model

- How do we get so many latent prints from Mr. X?
  - Create them digitally
  - Use a Distortion Model on the Known Print of Mr. X to digitally create many pseudo-latent prints
- How do we get so many latent prints from the Relevant Population?
  - A sufficiently large AFIS database can serve as the Relevant Population
  - Use the same Distortion Model on prints from AFIS to digitally create many pseudo-latent prints
  - Since we only care about the left side of the distribution, only distort the prints for the “closest” ~100 people
- How do we do all these comparisons?
  - Use the AFIS to measure the dissimilarity between the Latent Print and all the pseudo-latents



Note: Image is conceptual only and does not represent the actual method used to distort prints

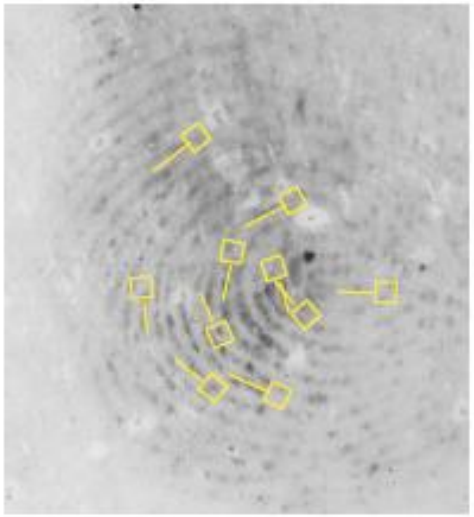
Does it Work?

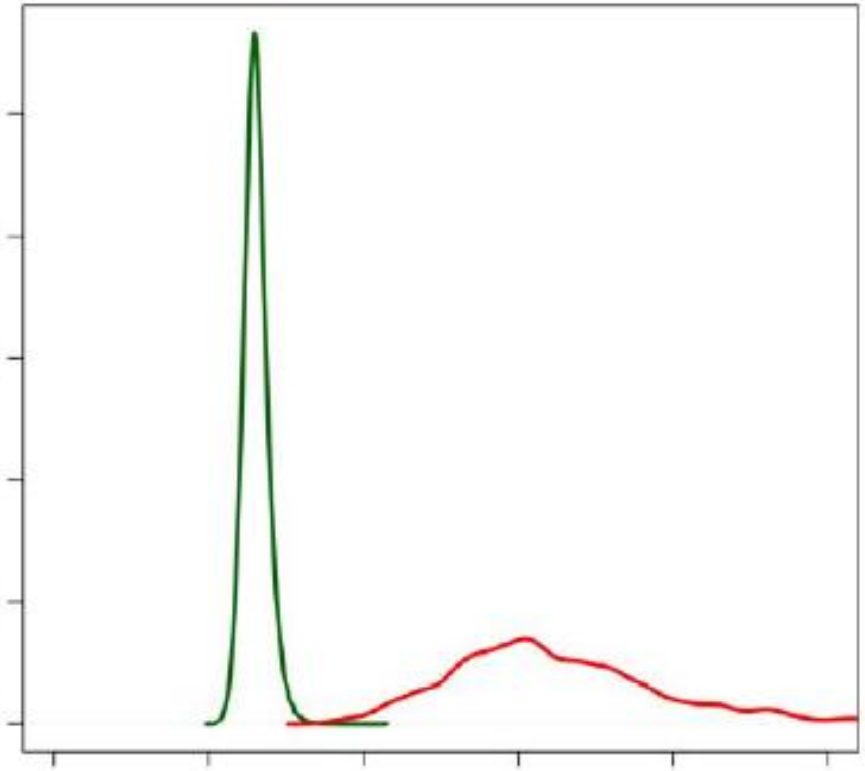
- Latent Print
  - 207 sample latent prints
  - 1837 minutiae configurations

# minutiae	6	7	8	9	10	11	12	13	14	15
# configurations	196	193	193	189	188	184	179	176	171	168

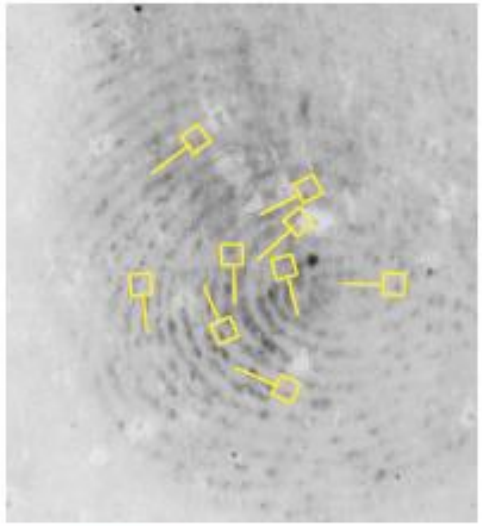
- Known Print
  - True Source finger – corresponding minutiae between each Latent Print and Known Print marked by a qualified LPE
  - Close Non-Match finger – highest scoring finger in AFIS database – corresponding minutiae marked by AFIS
- Distortion Model
  - 500k pseudo-latent prints from the TS or CNM
  - 500k pseudo-latent prints from 99 closest fingers in AFIS (not counting the CNM)
- Each LP/TS and LP/CNM sample was run 3 times

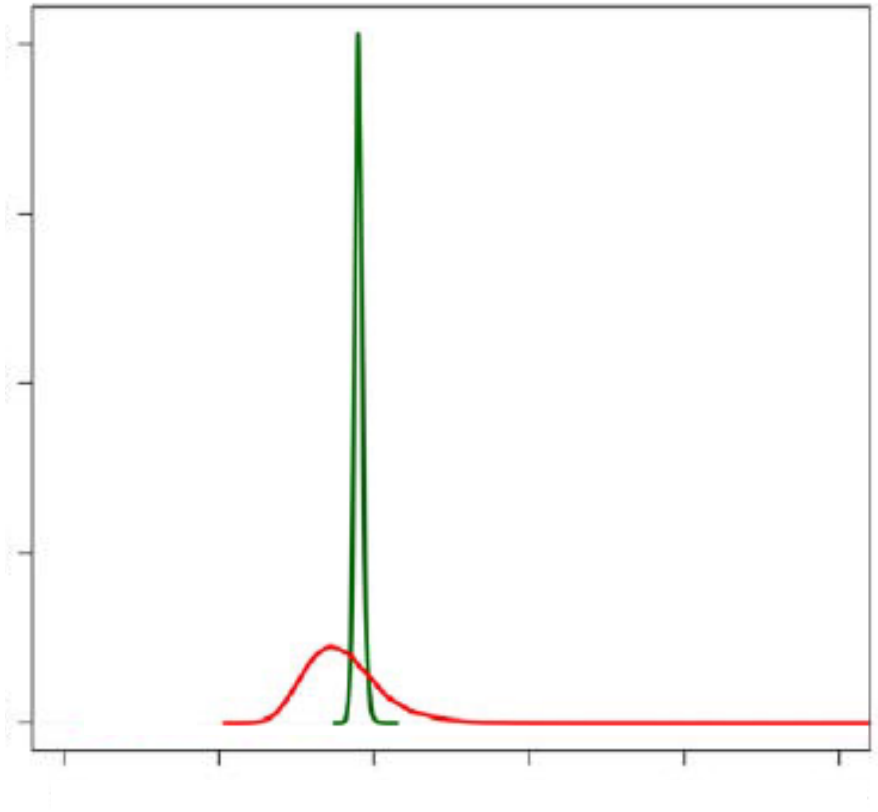
Latent Print and True Source

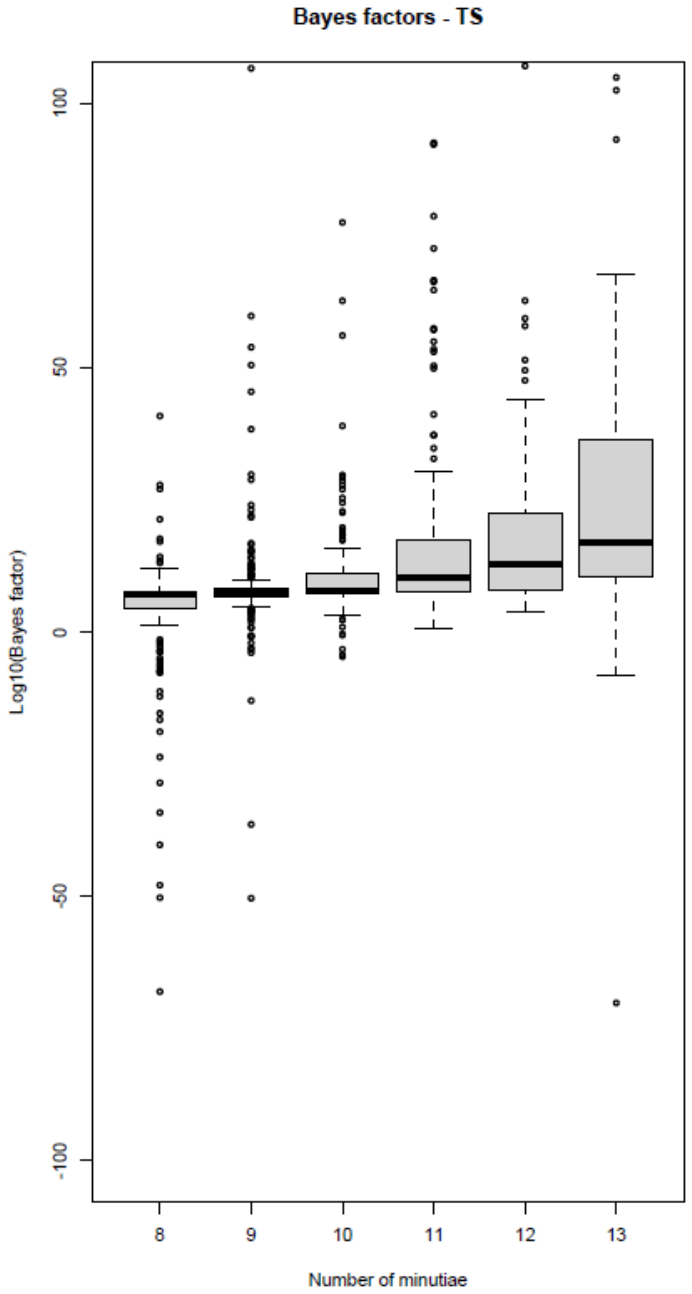


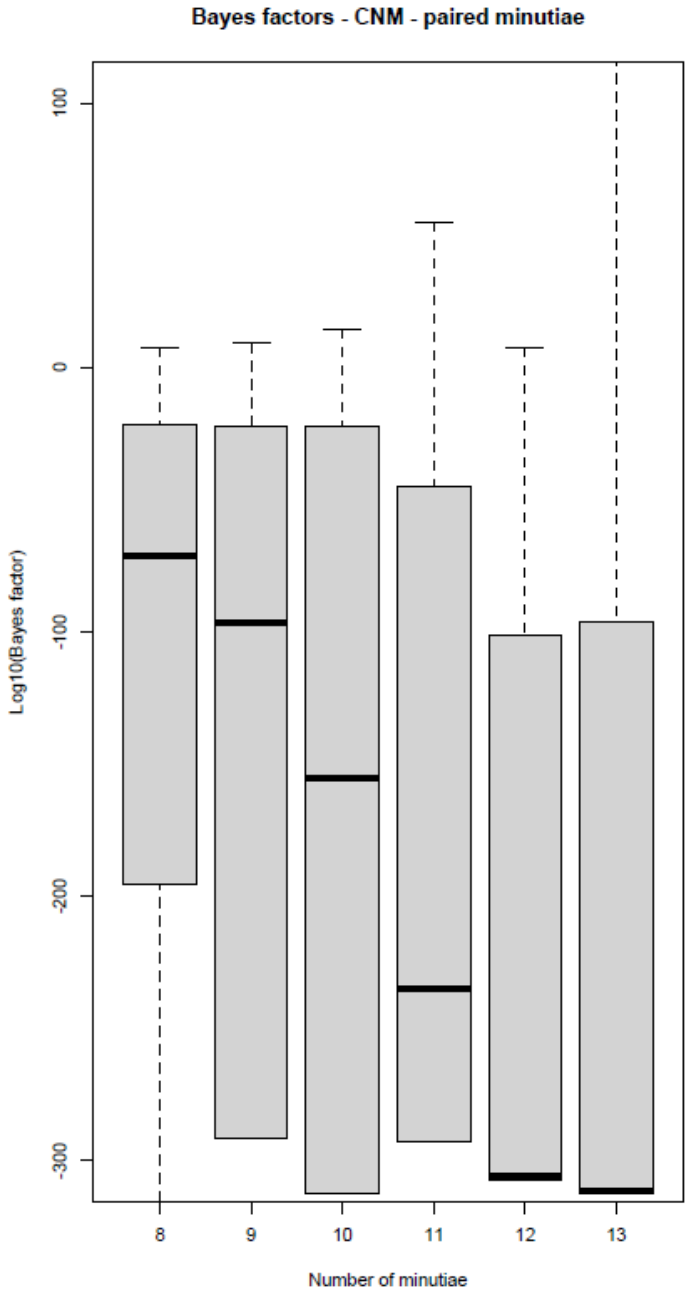


Latent Print and Close Non-Match





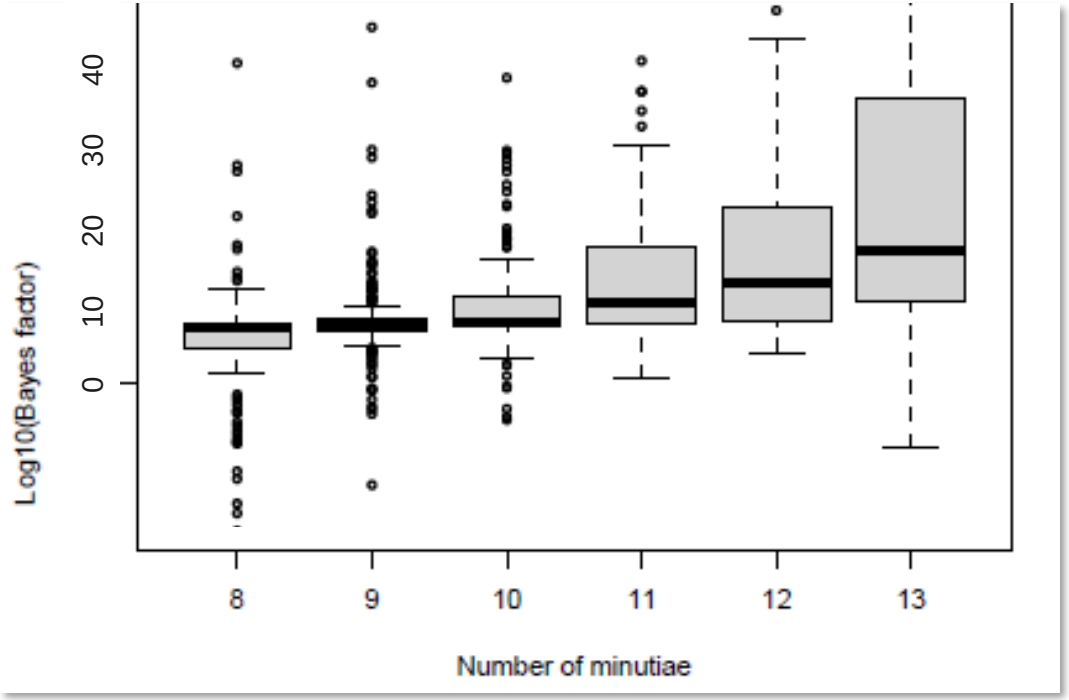

Bayes factors - TS
Log10(Bayes factor)
100
50
0
-50
-100
8
9
10
11
12
13
Number of minutiae

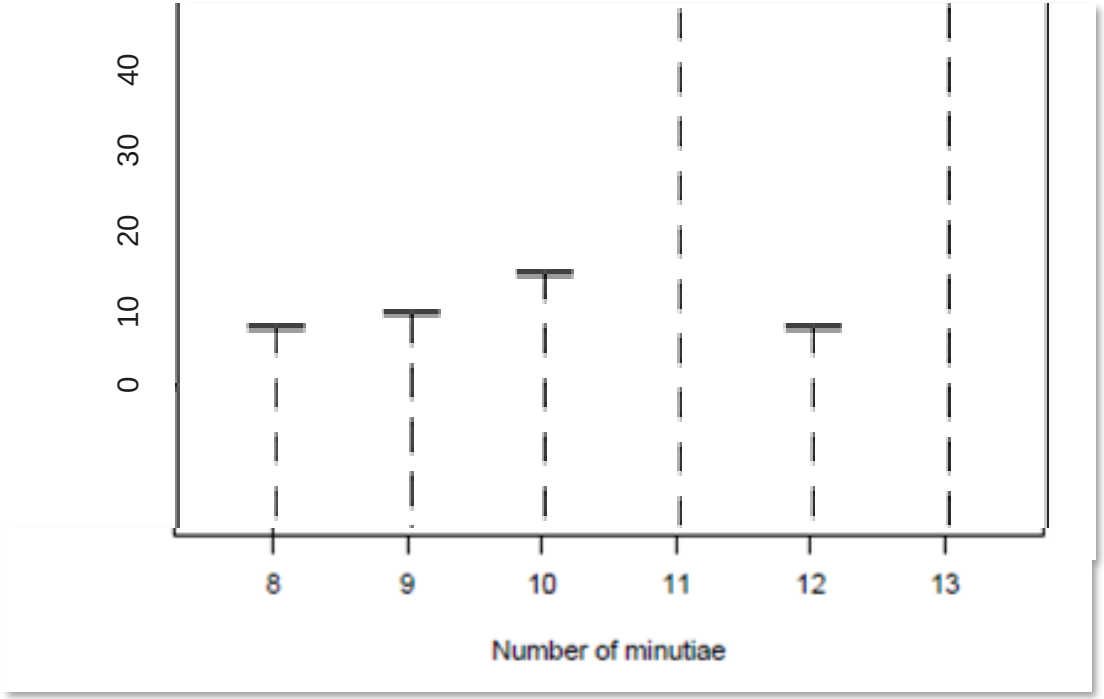

Bayes factors - CNM - paired minutiae
Log10(Bayes factor)
100
0
-100
-200
-300
8
9
10
11
12
13
Number of minutiae

IDEMIA

Test Results

- At minutiae configurations of **less than 8**, the metric was not discriminative enough
  - Could not distinguish results for TS and CNM tests
- At minutiae configurations **greater than 13**, the results were near infinity (for TS tests) or near zero (for CNM tests)





IDEMIA

Test Results

# minutiae	8	9	10	11	12	13
Rate of Misleading Evidence in favor of Prosecution (LR > 1 for a CNM test)	18%	17%	22%	16%	9%	17%
Rate of Misleading Evidence in favor of Defense (LR < 1 for a TS test)	28%	11%	6%	0%	0%	2%

- A Likelihood Ratio greater than 1 does not seem sufficient to support an Identification
- Further research is needed to set an appropriate threshold to minimize misleading evidence that would favor the Prosecution

What Next?

- Preliminary tests need to expand into a full validation study
  - IDEMIA does not have access to large datasets
  - Need to partner with law enforcement agencies to validate method
  - Explore topics related to non-mated minutiae and quality
  - Explore potential need for calibration
- Need to determine an appropriate threshold – above which would support an Identification conclusion
- Need to establish formal training on how to use this tool
- Need to establish formal training on how to describe this process and these results in a courtroom
- Need to establish agency policies on when to use this tool

References

- David R. Ashbaugh. *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology.* Boca Raton, CRC Press 1999.
- Francis Galton. *Finger Prints.* London, MacMillan and Co., 1892.
- Jessie Hendricks, Cedric Neumann, Christopher P. Saunders. "Quantification of the weight of fingerprint evidence using a ROC-based Approximate Bayesian Computation algorithm for model selection." Electronic Journal of Statistics, 15(1) 1228-1262 2021.
- Edmond Locard. *La Preuve Judiciaire par Les Empreintes Digitales.* Lyon, A. Rey, 1914.
- National Research Council. (2009) Strengthening Forensic Science in the United States: A Path Forward. Washington, DC: The National Academies Press.
- Cedric Neumann, Ian W. Evett, James Skerrett. “Quantifying the weight of evidence from a forensic fingerprint comparison: A new paradigm.” Journal of the Royal Statistical Society: Series A (Statistics in Society), 175(2), 2012, Pages 371–415.
- Cedric Neumann. “Defence against the modern arts: the curse of statistics: Part I—FRStat”, *Law, Probability and Risk*, Volume 19, Issue 1, March 2020, Pages 1–20.
- Cedric Neumann, Madeline Ausdemore. “Defence against the modern arts: the curse of statistics—Part II: ‘Score-based likelihood ratios’”, *Law, Probability and Risk*, Volume 19, Issue 1, March 2020, Pages 21–42.
- Cedric Neumann, Jessie Hendricks, Madeline Ausdemore. *Statistical Support for Conclusions in Fingerprint Examinations from*: **Handbook of Forensic Statistics**. Boca Raton, CRC Press, 2021.
- Cedric Neumann. (2019-2021). *Data-driven decision-making in fingerprint examination*. Unpublished reports.
- OSAC Friction Ridge Subcommittee. (2021) OSAC Research Needs Assessment Form: Friction Ridge Statistical Modeling. Organization of Scientific Area Committees for Forensic Science.
- PCAST. (2016) *Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods.* Washington, D.C.: Executive Office of the President’s Council of Advisors on Science and Technology
- Henry Swofford, et al. “A method for the statistical interpretation of friction ridge skin impression evidence: Method development and validation.” Forensic Science International, Volume 287, 2018, Pages 113-126.
- Melissa Taylor, et al. (2012) Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach, NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD

IDEMIA

QUESTIONS?

Eric Ray, CLPE
Technical Project Leader | IDEMIA

M. +1 657-230-0745
E. Eric.Ray@us.idemia.com

IDEMIA