The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	Coping with Close Non-Matches in Latent Print Comparison (re-)Training
Author(s):	Heidi Eldridge, Marco De Donno, Margaux Girod, Christophe Champod
Document Number:	305757
Date Received:	January 2023
Award Number:	2018-DU-BX-0227

National Institute of Justice – Office of Justice Programs

Research and Development in Forensic Science for Criminal Justice Purposes

Award # 2018-DU-BX-0227

Coping with Close Non-Matches in Latent Print Comparison (re-)Training

By Heidi Eldridge, Marco De Donno, Margaux Girod, Christophe Champod

Project Period: January 1, 2019 – September 30, 2022

Award Amount: $616,407

Final Research Report submitted September 30, 2022

From:
Heidi Eldridge, PhD (PI)
RTI International
3040 East Cornwallis Rd.
Research Triangle Park, NC 27709
Email: Heidi.Eldridge@icloud.com

To:
Greg Dutton, PhD
Physical Scientist / Program Manager
National Institute of Justice
U.S. Department of Justice
810 7th Street NW
Washington, DC 20531
Email: Gregory.Dutton@usdoj.gov

Summary of the Project

Major goals and objectives/Research Questions

During latent print comparison, a non-mated print will be colloquially qualified as a "close non-match" (CNM) to the mark when it shows a level of agreement that could have misled an examiner to erroneously conclude that it was from the same source as the mark whereas it is established that print and mark do not share the same source. Put another way, a CNM is a print from a different source that shows a higher level of correspondence with the features of a mark than a randomly selected print would.

These CNM prints are of crucial importance in developing the expertise of latent print examiners (LPE) because they constitute the worst-case scenario for a comparison between impressions originating from different sources.

Because of the nature of the algorithms used in Automated Fingerprint Identification Systems (AFISs) (which are designed to find the algorithmically closest matches contained in the database), combined with the ever-enlarging set of prints in their gallery, it is expected that examiners will increasingly face comparisons involving CNMs.

There is some evidence that, under controlled conditions, examiners can make wrong associations when a CNM is compared to a mark with low quality or quantity of features. In 2012, Langenburg et al. [1] were the first to report a black box study in which five of the cases involved CNMs obtained following the search of an AFIS database of approximatively 600 million fingerprint images. Among 158 experts, a false positive rate of 2.2% was reported. In a 2013 National Institute of Justice (NIJ)-funded study, Neumann et al. [2] counted 11 erroneous identifications (among 124 examiners viewing the image pair) for a single CNM case. Using the same images, Liu et al. [3] tested 29 experienced examiners and obtained 3 erroneous identifications. Referring to the systematic use of AFIS in China, they indicated (p. 36): "As the database size is increasing, the examiners' experience alone is no longer sufficient to deal with CNM prints." Koehler and Liu [4] conducted a much larger study providing examiners from 125 Chinese agencies with mandatory proficiency tests that included two CNMs. One of these CNMs was the same pairing used in the earlier Neumann et al. and Liu et al. studies. Their findings included a shocking 15.9% false positive error rate for the reused CNM pair and an even higher 28.1% false positive error rate for a novel CNM pair.

The issue facing the fingerprint community is that it does not have any repository of known cases including CNMs from which comparison exercises can be drawn. This is largely because CNMs are so time-consuming to locate. Over time, anecdotal cases have been shared in the community [5], but without any systematic organization to the data or means to search them available to law enforcement agencies (LEA). Additionally, these cases were submitted by examiners from casework, thus the ground truth state of the image pairs was unknown. Finally, although examiners are aware of the potential risks associated with CNMs, it is rare for any of their training or proficiency test materials to incorporate them. The end result is that essentially, examiners have no mechanism to find their limits or to identify red flags to predict when CNMs may lead them into an error. It is unclear at this point if examiners can identify areas of friction ridge skin (FRS) that are more prone to CNMs than others.

To address this problem, we undertook to develop such a sanitized close non-match database for research and training that will be shared with trusted members of the LEA community. We call this database the International Close Non-Match Library (ICNML). Thanks to partners at numerous Unites States and international LEAs, we collected known ground-truth marks and prints from 100 donors and searched marks through multiple AFISs both in the US and internationally to obtain CNMs. As a research question, we explored whether specific groups of criteria, or red flags, could be used to yield a higher number of CNMs when searched in large AFIS databases. If so, these criteria could be developed into training for LPEs on when to exercise additional caution around identification decisions that may be more likely to lie in a danger zone for a CNM.

Research design and methods, analytical and data analysis techniques

International LEAsourcing

To populate the ICNML, an international team of LEA partners was assembled. This was done for two main reasons: (1) to distribute the workload among many partners because it was too much for only a few operational laboratories to complete while maintaining their workloads; and (2) to draw from multiple jurisdictions, both because different areas have different rules about AFIS searching and sharing of results and because we wanted the ability to search each selected area in more than one AFIS to increase our chances of locating CNMs.

We call this "LEAsourcing" as it is a kind of distributed-labor model similar to crowdsourcing, but relying solely on LEA partners.

We defined several different roles for this project, each completing a distinct set of tasks. These roles were called "teams" and are briefly described below. A partner LEA agency could be assigned to one or more teams, depending upon their availability to do the work, the skillsets of their particular staff, and the prevailing rules of their jurisdiction or agreement of their legal teams.

Detection Team

The main task of the Detection Team was to recruit donors and collect marks and known prints from those donors. All donors were provided with informed consent, which was approved by RTI's Institutional Review Board (IRB) prior to any collection.

The first task of the Detection Team was to produce a single, pristine set of finger and palm exemplars from each donor who completed the informed consent and upload them into ICNML. These pristine exemplars were then passed to the Selection Team to select 5 areas of interest.

Once the areas of interest were determined, the Detection Team was responsible for going back to their donors and creating 10 marks from each of the 5 selected areas of interest (for a total of 50 marks from selected areas). These marks were made and collected under observation so that their true source was established. They were intended to be from a broad variety of matrices on a broad variety of substrates and developed using a broad variety of development methods. They were also intended to represent a wide range of quality, but always to be suitable for comparison (if sometimes only barely).

In addition to the 50 marks from selected areas, the Detection Team was responsible for selecting an additional 5 marks from each donor that were *not* from selected areas. These marks were referred to as "incidental" marks and came from the normal handling of objects while the selected area marks were being produced (e.g., if a glass was held in order to capture the selected area of an index finger, an incidental mark might have been produced and selected from one of the other fingers or the thumb that also touched the glass as it was being held normally).

Finally, the Detection Team was responsible for generating two additional full sets of finger and palm exemplars – one of medium quality, and one of poor quality. All marks and prints were uploaded into ICNML without any identifying information beyond the impression images themselves and were associated with an anonymized record for that donor.

Selection Team

The Selection Team was comprised of 10 handpicked experts who had experience in locating CNMs. Near the beginning of the project, an in-person workshop was held during which the 10 Selection Team experts and representatives from each of the partner agencies met to discuss the risk factors for a CNM. Through examples, discussion, and brainstorming sessions, we developed a list of criteria that the team believed were more likely to result in CNMs during an AFIS search than areas that did not conform to these criteria. From this list of criteria, we developed rules for AFIS searching to balance our sample types and locate CNMs.

We determined the following distribution for the five areas that would be selected from the exemplars of each donor:

- 1 impression from the palm
- 1 delta area from a finger
- 1 area along or near type lines
- 1 pattern focus area
- 1 selector's choice (preferably from a finger)

We additionally stipulated that if a particular donor did not exhibit a suitable area in the pattern focus area category, the selector may instead make a second "selector's choice" selection.

For each of the three broad search categories (fingers, pattern focus areas, and palms), we further defined specific areas or criteria that were preferred for selecting an area as they fit into our criteria of being more likely to yield CNMs. These are briefly outlined as follows:

<u>Fingers</u>

Deltas

In particular, we sought deltas with many small minutiae near their center, or deltas with many ridge endings flowing in the same direction.

Common ridge patterns

We sought to select target groups for searching in areas where particular ridge cluster types are known to be commonly present and therefore more likely to be repeated, such as:

- Groups of ridge endings away from cores, such as in the area outside the delta, on the side of the distal phalanx
- Groups of features along type lines or between the type lines and distal crease, particularly if there was an ambiguous small feature (such as an enclosure or handshake) nearby.
- Small enclosures or switchbacks below the cores of large whorls
- Groups of ridge endings in the pattern force outflow area of a loop or delta
- Loop cores with a single rod at the center and clusters of minutiae to the side of the core, but not above it
- Plain whorl cores with a single short ridge at their center
- A series of bifurcations adjacent to one another next to the core of a loop

<u>Pattern Focus Areas</u>

There are particular pattern types in fingers that we concluded were more likely to share common features around their core areas and we designated these as “pattern focus areas”. They included:

- Low count loops
- Arches

- Tented arches
- Low count central pocket loops

Palms

There were four areas on palms that we identified as having a high probability of common features between individuals and we selected these preferentially when they met our overall considerations:

- Deltas (following the same rules as outlined for finger deltas, above)
- The area below high carpal deltas, if that area did not have many creases
- Any place on the palm where ridges come together in a funnel-like flow (not just the "funnel" area defined by Ron Smith, but any place where ridge systems are converging and many ridge endings are being forced)
- Areas of mostly straight, parallel ridges without distinguishing anchors nearby

Overall Considerations

The following considerations were preferred for selected areas:

- Each selected cluster of minutiae should contain between 6 and 10 minutiae at the selector's discretion
- When possible, a cluster of minutiae should be chosen that has a small, ambiguous feature nearby[1]
- Any area where there were many ridge endings in the same direction was preferred

[1] The reasoning for this can best be exemplified by the Mayfield error, where a small feature near the bottom of the impression could have been interpreted as either an enclosure or a dot/very short ridge. This feature stood out and could reasonably be given a lot of weight in comparison, yet its ambiguous nature means that it could easily mislead an examiner into incorrectly interpreting it and then overweighting it, leading to a false positive error in a CNM.

- No fingertips or extreme sides of fingers were selected, because they are typically not captured in exemplars recorded in AFIS

Once the first set of pristine marks from each donor was uploaded into ICNML, the Selection Team was responsible for reviewing all the finger and palm impressions that had been submitted and selecting 5 areas for each donor that they felt were likely to yield a CNM in an AFIS search according to the agreed-upon criteria. Because 10 experts viewed the exemplars from each donor, they were not always in agreement in which 5 areas they selected. Final, aggregate selections were made by the Coordination Team.

AFIS Team

LEA partners were invited to be part of the AFIS team if they were able, through local law and regulations, and by the consent of their legal departments, to sign a data use agreement with RTI International that stated that images located through AFIS searches could be shared with the ICNML, with the understanding that these would be accessed by trusted LEA partners for use in training, testing, and research. No personally identifiable information (PII) other than the images themselves was collected or stored in the ICNML; images of the finger and palm exemplars of selected candidates were uploaded in .NIST or .tiff format without any associated metadata such as name, data of birth, identification number, or sex.

After all marks (50 from selected areas and 5 incidentals) for each donor had been uploaded into ICNML, the AFIS Team was responsible for searching them in AFIS to locate CNMs. Impressions to be searched were distributed to LEA partners such that no laboratory searched impressions that came from its own donors.

Each AFIS search yielded two candidates whose exemplar impressions were uploaded into ICNML. The first candidate was always the number 1 candidate returned by the AFIS. If the examiner conducting the AFIS search felt that this candidate represented a good CNM, the second candidate on the list was also uploaded. If the examiner did not feel the number 1 candidate was a good CNM, they would then review the candidate list and select the candidate that they felt was the closest CNM available and that record was uploaded. For each candidate uploaded, the AFIS examiner indicated whether or not they personally felt that the candidate represented a good CNM.

Two different types of impressions were searched through AFIS: impressions from selected areas and incidental marks. Impressions from selected areas were either the marks that were deliberately made from selected areas of interest or prints from the same areas. Each selected area was searched through two different AFISs at two different LEA partner laboratories; one laboratory searched a mark made from the area of interest whereas the other laboratory searched a print of the same area. The specific minutiae to be marked and searched in AFIS were pre-determined by the Coordination Team as the aggregate minutiae marked by the Selection Team.

Incidental marks were only searched through a single AFIS. Since only one impression existed for each incidental area, all 5 incidental areas were searched. For the incidental marks, no set minutiae or criteria were employed; the AFIS examiner was free to mark and search the impression according to their usual policies and preferences.

Coordination Team

The Coordination Team was comprised of the researchers (article authors) and was responsible for oversight of the other teams and for key decision points.

The Coordination Team was responsible for final selection of areas and minutiae to be used in AFIS searching. The Coordination Team reviewed the selections made by the Selection

Team and made a tally of the “votes” that were received for different areas. The areas that were chosen by the largest number of Selection Team members were chosen as the 5 areas to be sent back to the Detection Team to create marks of, keeping in mind the desired distribution of 5 different areas (e.g., if two delta areas received 5 votes each and an area along the type lines received only 3 votes, the type line area and one of the delta areas would be selected so that both delta and type line areas would be represented. In this hypothetical situation, the Coordination Team member would decide which of the two tied delta areas would be selected based upon either the number of minutiae marked in common by the Selection Team, or the interest of the area/minutiae and the likelihood—in the Coordination Team’s opinion—that this area would yield interesting results).

Once an area was selected, the Coordination Team would review the specific minutiae that were selected by the Selection Team and make a final selection of minutiae to be entered by the AFIS Team. This final selection was based on a combination of number of “votes” for the minutiae, the placement of the minutiae, and a target total. Most selected areas had between 6 and 10 minutiae selected.

Specifications and contents of the ICNML

The ICNML was designed to contain a wide range of known source impressions including same source and different sources exemplars. We populated the ICNML with approximately 5,500 marks from 1,000 distinct areas of the friction ridge skin of 100 different donors. For each donor, 10 marks were produced from each of 5 areas that were selected by the Selection Team (50 marks per donor). In addition, for each donor, 5 more incidental marks that were created during the normal handling of objects were collected by the Detection Team. This resulted in a total of 55 marks collected per donor, or a total of 5,500 marks.

Each mark is associated with a minimum of 3 known same source prints. Full sets of exemplars (fingers and palms) were taken from each donor at low-, medium-, and high-quality levels. This allows for comparison exercises of variable difficulty to be constructed by selecting exemplars of the desired difficulty level.

Each mark is also associated with two or four sets of known different source prints. Each mark from a selected area was searched through two different AFISs and two candidates were returned from each search (4 total sets of exemplars) whereas each mark from an incidental area was searched through a single AFIS and two candidates returned (2 total sets of exemplars). This resulted in an ICNML-wide total of approximately 3,000 known different source prints including the closest CNMs we could locate for each distinct area of friction ridge skin.

Security

The security of the data collected for the ICNML was taken very seriously. ICNML has been elaborated adopting a privacy-by-design approach to provide all guarantees to donors of the appropriate level of security of their personal data (including personal information and biometric information). The global architecture of ICNML is based on encrypted data (AES256) with data encryption keys (DEK) residing with the donors only. None of the users or administrators of ICNML have access to the personal data associated with the donors. These data are hashed with their DEK. Donors received their own credentials on ICNML and maintain the right to remove their entries from ICNML, at any time. In other words all images that can be seen on ICNML are encrypted with a DEK in possession of the donors. The full architecture of ICNML has been developed using open-source libraries and is available for review at https://esc-md-git.unil.ch/ICNML.

No personally identifiable information (PII) was collected from donors or candidates located through AFIS search, other than the impression images themselves. Login requires time-based one-time password (OTP) two-factor authentication for anyone who accesses the database in any role (physical security keys were required for admin accounts). No agency was or will be granted access to ICNML unless they have been vetted and live in a country that has contributed donor data to the ICNML and have signed a Data Use Agreement limiting how they may use the data. All downloaded data packets have digital watermarks that will allow the researchers to trace any improperly used or distributed data. And finally, donors were given the ability to remove their own data from the database at any time, for any reason, and without notifying anyone of their decision.

Longevity

ICNML has been designed to be of value to the international forensic science community for years to come. Although donors have the ability to remove their own data from ICNML, which could lead to a loss of data over time, ICNML has been designed to also allow for new donors to be *added* over time; hopefully resulting in a net gain in records stored within the database.

As noted above, and in contrast to some existing fingerprint research databases that are publicly available, ICNML has restricted access to only trusted law enforcement partners. The purpose of this restriction is twofold. First, it allows the researchers to vet LEAs that want access to the database to ensure that they are from countries that safeguard the human rights of their citizens and that they have good reputations for data integrity. This minimizes the risk of granting access to agencies that might not take the privacy and integrity of the data seriously and thereby allows all countries and agencies who have donated images to the database to feel more secure that this information will not be abused.

Second, it allows the researchers to offer an incentive to countries and agencies to contribute additional data to the database. Only agencies within a country that has donated data to the ICNML are eligible to apply for access to the database. Therefore, as more countries become interested in the ICNML and want to gain access, they will need to contribute donors, which will help to increase the number of records available for everyone's use. By contributing data from their own citizens to the database, each country has a vested interest in ensuring that those data are protected and properly used.

Through this two-fold mechanism (requiring countries to donate user data to gain eligibility to apply for access and vetting individual agencies within those countries before granting access), ICNML will be able to maintain its size or grow larger and keep the data it contains safe for years to come.

Selection of experimental trial images

To establish whether the CNM criteria developed by the expert working group were effective in predicting areas that were more likely to yield CNM pairings when entered in a large AFIS, we devised an experiment to measure the frequency of false positive errors made when comparing pairs that were found in AFIS using the CNM criteria compared to those made when comparing pairs that were found in AFIS using incidental marks that were not selected based on any criteria.

To test the difference in number of CNMs generated by targeted marks versus incidental marks, two conditions were needed—different sources comparisons of target marks and different sources comparisons of incidental marks. However, we did not want participants to have an assumption that all (or even most) of the comparisons they were asked to do were different sources pairings. Thus, we introduced a third condition, which was same source pairings, as a guard against bias.

Each participant in the study was asked to complete 12 comparison trials. Unbeknownst to them, these 12 trials were comprised of 6 same source pairs, 3 different sources pairs from target cases, and 3 different sources pairs from incidental cases. The order of the 12 trials was randomized on the list of assigned cases, and participants had the ability to work the cases in any order they chose.

To control for the difficulty of the mark as an influencing factor on the false positive error rate, we used each mark in two different comparison trials – once in a same source pairing and once in a different sources pairing. Thus, if a particular mark had a very high false positive error rate, but participants also struggled to correctly identify it in the same source trials, we would be alerted that the mark was difficult and could be influencing error rates.

To ensure that no single participant saw the same mark twice (i.e., in a same source trial and in a different-sources trial), we divided the participants into 3 groups. They were assigned rotationally in the order they enrolled in the study (i.e., the first person to sign up was put in Group 1, the second person was in Group 2, the third in Group 3, the fourth in Group 1, and so on). The groups were assigned cases as illustrated in Figure 1 so that the marks from each different-sources condition moved as a block to a different group to be the same source trials. For example, Case 1_C (a different sources trial for Group 1) used the same mark as Case 2_E (a same source trial for Group 2).

	Group 1	Group 2	Group 3
Diff. Sources Target Cases	Case 1_C	Case 2_C	Case 3_C
	Case 1_I	Case 2_I	Case 3_I
	Case 1_A	Case 2_A	Case 3_A
Diff. Sources Incidental Cases	Case 1_F	Case 2_F	Case 3_F
	Case 1_K	Case 2_K	Case 3_K
	Case 1_B	Case 2_B	Case 3_B
Same Source Cases	Case 1_E	Case 2_E	Case 3_E
	Case 1_H	Case 2_H	Case 3_H
	Case 1_L	Case 2_L	Case 3_L
	Case 1_D	Case 2_D	Case 3_D
	Case 1_J	Case 2_J	Case 3_J
	Case 1_G	Case 2_G	Case 3_G

Figure 1 Assignments of cases from three different conditions to the three groups of participants. Cases in corresponding positions in the same color block used the same mark. For example, the same mark was used in Case 1_C, a different sources target case given to Group 1, as was used in Case 2_E, a same source case given to Group 2.

Cases were selected for inclusion in the experiment based upon the different-sources pairings. Once the 18 different-sources cases had been selected (9 target marks and 9 incidental marks), the same 18 marks were paired with one of their same source exemplars to form the same source trials.

To avoid an unconscious bias on the part of the researchers by which the hardest CNMs from the target marks and the easiest different-sources pairings from the incidental marks might be chosen to support our hypothesis that more CNMs would be found in the pool of target marks, we randomized the selection process to the extent possible.

First, we created a pool of putative CNM pairings from which to draw. To do this, we relied upon the opinions of the AFIS team members who did the AFIS searching and uploaded the closest CNMs they could locate for each mark. At the time of uploading each AFIS result, AFIS team members were asked whether, in their opinion, the exemplar being uploaded was a ‘good CNM’ to the searched mark.

When creating our pool of CNMs, we first created a list of 467 cases selected randomly from the ICNML. We then proceeded through this list, checking the response of the AFIS team member to the ‘good CNM’ question for each pair. If the response was “no,” the pair was not included in the pool. If the response was “yes,” the pair was included in the corresponding pool (target marks or incidental marks).

Once the pools were populated, the authors met as a group and reviewed each case in the pools, classifying them as green, yellow, or red where green were good CNM pairs by the consensus of the authors, yellow were mid-range in difficulty, and red were not considered to be challenging CNM pairs.

Our intent was to then randomly select 9 green cases from each pool, representing the closest CNMs that were found in each pool by a random sampling process. However, as it happened, there were exactly 9 green cases in the incidental marks pool, so all 9 of these were selected for use in the experiment. The targeted marks pool contained 26 green pairs, so 9 cases were randomly selected from among these 26 for inclusion in the experiment.

Once the different sources pairs had been chosen, the mark images were cropped to remove obvious differences between the marks and different source prints (such as different pattern types, different ridge counts, missing minutiae that were clearly present in one image, etc) to prevent obvious exclusions from being easily made. The same cropped mark images were used in the same source trials, so that there were no differences in visual cues between the same source and different-sources trials (i.e., it wasn’t obvious which were the same source trials due to them not being cropped; *all* marks were cropped and the pairs of marks that were the same between two trials were cropped the same way).

Experimental participants and process

LPEs who were signed off for independent research were recruited to take part in the experiment using an email list maintained by the primary author of LPEs who are interested in participating her research studies. Invitations to participate were sent to the list via a confidential liaison who managed enrollment in the study such that none of the researchers were aware of the identities of any of the participants, each of whom was assigned an anonymous username. No demographic information was collected from the participants because this was not an error rate study and was not measuring their accuracy or reliability for the purpose of establishing discipline error rates, thus their age, experience, and other demographic information was irrelevant. All participants were provided with informed consent, which they agreed to before beginning the experiment, and which had been reviewed and approved by RTI's Institutional Review Board (IRB).

Comparisons were made onscreen using the online interface PiAnoS, which was developed by the University of Lausanne (https://ips-labs.unil.ch/doc/). PiAnoS allows participants to sign in using a unique account, view the cases assigned to them, and annotate each mark on its own before proceeding to a side-by-side comparison screen were minutiae in common could also be annotated. Participants were aware that this was a CNM study, that all marks were cropped, and that they would be presented with both same source and different-sources trials (although not the proportion of each).

The only response participants were required to give for each trial was their final conclusion regarding the source of the pair of impressions. They had three conclusion options available:

- Same source
- Different sources (CNM)
- Different sources (not a CNM)

Participants were instructed that no inconclusive option was available to them and that this was a forced-decision study design, but were reassured that this was not an error rate study and that they were not being judged on their accuracy. This design decision was made because the participants were aware this was a CNM study and if ‘inconclusive’ were an option, they might have just chosen it on every difficult trial to avoid being tricked by the CNM. Since we wanted them to be potentially influenced by the CNMs so that we could measure the difference (if any) in the rates of false positive errors in the target mark trials versus the incidental mark trials, we had to take this option away from them.

Participants were further instructed that if they determined a pair of images were from different sources they should select the appropriate response to indicate whether, in their opinion, the comparison was a CNM, or was an easy exclusion that did not represent a CNM. These qualitative data were collected as another way to measure which trials represented CNMs, particularly if the quantitative data should prove to be inconclusive (e.g., if no false identifications were made in any of the trials).

Participants were given approximately three weeks to complete their 12 trials and were free to save their work in progress and come back to it at their leisure.

Analytical and Data Analysis Techniques

Descriptive representations were made of the conclusions received in the experiment and are reproduced in the results section below. These analyses were carried out in R [6] in the RStudio IDE version 2022.7.1.554 [7] and taking advantage of the *tidyverse* set of libraries [8].

Expected Applicability of the research

The establishment of a database containing known same source and different sources impressions will strengthen the practice by providing law enforcement agencies with a tool to create training, testing, and research packets. This would be true whether or not good CNMs were found. However, any CNMs that have been located through this process will only enhance the value of the database as they will provide for more challenging exercises and as we learn more about their attributes, help to create training to reduce the numbers of erroneous identifications effected due to CNMs.

Participants and other collaborating organizations

Participants were recruited through email lists maintained by the primary researcher of LPEs who are interested in participating in research. Latent print examiners who were signed off for independent casework were invited to participate. Participants enrolled voluntarily in the study and their identities were held confidential using anonymous usernames. All contact with participants was done through a confidential liaison so the researchers were never aware of participants' identities. All possible link between usernames and true identities has been destroyed. All participants read and signed an informed consent statement that had been reviewed and approved by RTI's Institutional Review Board (IRB) prior to participation.

Several domestic and international LEAs and latent print experts collaborated on this project by participating on one or more of the teams helping to populate the database. These included: Université de Lausanne (Switzerland), National Forensic Science Institute (Romania), National Forensic Centre (Sweden), Division of Identification and Forensic Science (Israel), East Midland Special Operations Unit – Forensic Science (UK), Metropolitan Police Servies (UK), Sûreté de Québec (Canada), Royal Canadian Mounted Police (Canada), Elite Forensic Services, Evolve Forensics, HJS Consulting, Los Angeles Sheriff's Department, Arizona Department of Public Safety, Houston Forensic Science Center, Minnesota Bureau of Criminal Apprehension, and Tom Wortman.

Outcomes

Activities/Accomplishments

We undertook a number of activities in the completion of this project. Chief among them were:

- Hold an expert working group meeting to develop criteria for searching for CNMs
- Collect targeted latent impressions of known source along with 3 sets of exemplars each from approximately 100 donors
- Select areas from exemplars likely to generate CNMs according to criteria developed by expert working group
- Search a sub-set of marks and prints through multiple domestic and international AFISs to locate the closest CNMs we could
- Create and populate the ICNML database to house all generated impressions in a secure environment
- Recruit participants to take part in a small study to test the efficacy of our criteria and provide participants with instructions and informed consent
- Update PiAnoS to meet the needs of the project
- Select images for use in the study and populate case lists for participants
- Manage data collection, including sending responses to anonymized participant queries
- Extract and clean data
- Graphically summarize data

- Prepare and deliver dissemination materials, including manuscript for submission to a peer-reviewed scientific journal, and conference and educational meeting presentations

Results, Findings, and Limitations

The ICNML database was successfully populated with known impressions from nearly 100 donors. At the time of writing, CNMs are still being added to the database as a result of continued AFIS searching. This database represents the first effort of its kind to produce a large database of known same source and different-sounces impressions that includes the closest CNMs that could be located from multiple international AFIS searches and that will be made available to trusted law enforcement partners for training, testing, and research.

The experimental phase to test the effectiveness of the developed CNM criteria yielded data from between 40 and 47 LPEs viewing each of our 12 trials. The summary data are presented in Figure 2.

From these data, it should be noted that although there were many opinions of “CNM” among the different sources trials, there was no consensus on trials that all, or even most, participants thought were CNMs. It should also be noted that although very few false positive conclusions were made, more of them were made in the incidental cases than in the target cases.

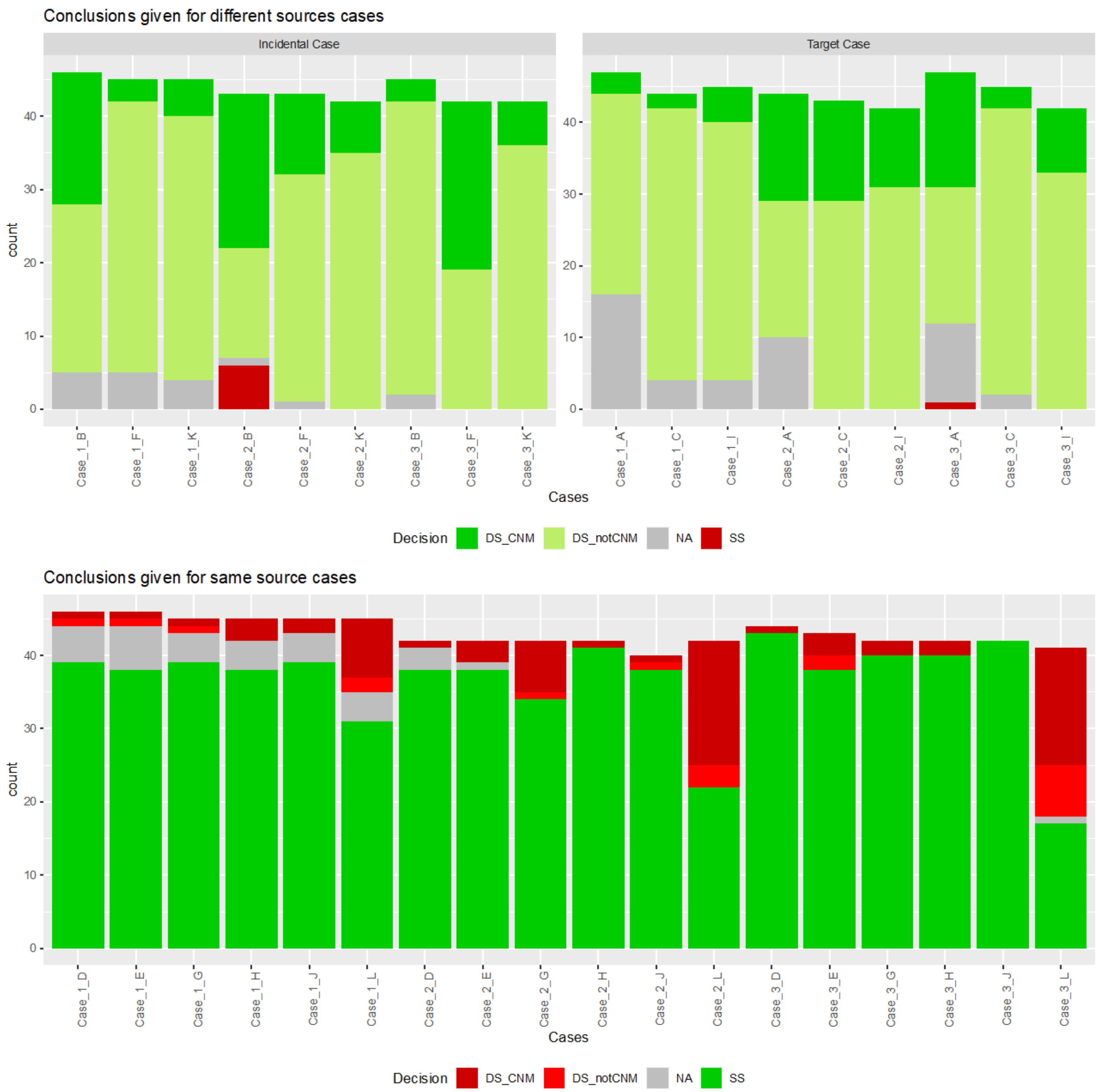

Figure 2. Conclusions given in the experimental phase showing how many false identifications were made in Target cases versus Incidental cases as well as how many different sources trials in each condition participants thought were CNMs.

At face value, this would seem to indicate that the incidental cases were more likely than the target cases to produce CNMs, suggesting that the CNM criteria developed by the expert working group were not, in fact, successful in predicting areas that were more likely to yield

CNMs. However, additional factors need to be taken into account and represent limitations of the current study.

First, the experimental participants did know that this was a CNM study. Although we tried to counter this knowledge by including 50% of the trials that were same source trials, the participants were nonetheless likely on guard against CNMs, which could have artificially lowered the incidence of false positives across the board.

Second, the image pairs for the experimental phase were selected through a random process. Although this was done to reduce researcher bias effects, it also means that the target group condition was not necessarily populated with the closest CNMs available through searching using the developed criteria. In fact, the researchers are currently unaware of how many close CNMs the ICNML contains, nor whether these were generated through Target searches or Incidental searches. Now that the impressions have been collected, a good extension project will be to comb through the AFIS candidates that were generated and flag those that appear to be the best CNMs. These could then be blindly presented to a number of practicing LPEs to test the CNMs' ability to fool LPEs into erroneous identifications and a tally made of the relative distribution between the Target and Incidental groups.

Some of these data can be collected as users begin to create packets in ICNML and report back to the researchers when they encounter what they believe to be good CNMs through testing and training. The database will only become stronger and richer over time as new records are added and we learn more about its contents.

Artifacts

List of products

The following products have been created as a result of this research:

- ICNML – the International Close Non-Match Library, which will be made available to trusted law enforcement partners for use in training, testing, and research. The full code of the ICNML infrastructure is available at https://esc-md-git.unil.ch/ICNML.
- Scientific article manuscript with working title “Developing an International Close Non-Match Library (ICNML) for latent print testing, training, and research” in final preparation for submission to a peer-reviewed journal.
- Conference presentations of results at the following conferences and educational meetings:
    - 2021 International Association of Identification (IAI) Educational Conference
    - 2022 American Academy of Forensic Sciences (AAFS) Educational Conference
    - 2022 IAI Educational Conference

Data sets generated

The data collected from participants in this study include the contents of the ICNML and the comparison results of the experimental phase. All raw data have been archived by RTI and UNIL.

Dissemination activities

The results of this research have been disseminated in multiple ways, which have been listed above, under “List of products.” The results have been presented at numerous educational conferences; discussed thoroughly in manuscript that will be submitted for publication; and the database itself will be made available to trusted partners in the international law enforcement community.

Bibliography

1. G. Langenburg, C. Champod, T. Genessay, Informing the judgments of fingerprint analysts using quality metric and statistical assessment tools, Forensic Sci. Int., 219 (2012) 183-198.
2. C. Neumann, C. Champod, M. Yoo, T. Genessay, G. Langenburg, Improving the Understanding and the Reliability of the Concept of "Sufficiency" in Friction Ridge Examination, National Institute of Justice, Washington DC, 2013.
3. S. Liu, C. Champod, J. Wu, Y. Luo, Study on Accuracy of Judgments by Chinese Fingerprint Examiners, J. Forensic Sci. Med., 1 (2015) 33-37.
4. J. Koehler, S. Liu, Fingerprint error rate on close non-matches, J. For. Sci., 66(1): 129-134. doi: https://doi.org/10.1111/1556-4029.14580
5. Close calls. Available at https://web.archive.org/web/20150912071214fw_/http:/www.clpex.com/CloseCalls/CloseCalls.htm
6. R Core Team (2021). R: A language and environment for statistical, R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
7. Studio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.
8. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L,
   Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse."
   Journal of Open Source Software, *4*(43), 1686. doi: 10.21105/joss.01686 (URL: https://doi.org/10.21105/joss.01686).

.