Journal of Forensic and Legal Medicine 83 (2021) 102254

Contents lists available at ScienceDirect

Journal of Forensic and Legal Medicine

journal homepage: www.elsevier.com/locate/yjflm

Research Paper

A Swiss collaborative exercise for Disaster Victim Identification (DVI): Covering situations with different levels of complexity

Christian Gehrig [a,*], Séverine Delémont [a], Jennifer Comte [a], Tacha Hicks [a,b], Patrick Basset [a], Frédéric Grosjean [a], Daniel Dion [c], Christian Cossu [d], Michel Bottinelli [e], Mirco Hecht [f], Andrea Sulzer [g], Pamela Voegeli [g], Vincent Castella [a]

[a] University Center of Legal Medicine, Lausanne, Geneva, Switzerland
[b] Fondation pour la Formation Continue Universitaire Lausannoise (UNIL-EPFL), Switzerland
[c] Institut für Rechtsmedizin, Basel, Switzerland
[d] Institut für Rechtsmedizin, St. Gallen, Switzerland
[e] Laboratorio di Diagnostica Molecolare, Lugano, Switzerland
[f] Institut für Rechtsmedizin, Bern, Switzerland
[g] Institut für Rechtsmedizin, Zürich, Switzerland

ARTICLE INFO

Keywords:
Disaster victim identification
DVI
Simulated DVI exercise
Familias
CODIS
LR

ABSTRACT

The identification of victims of a disaster (DVI) requires the collaboration of different specialists. Within a DVI context, DNA analyses often play an important role. Consequently, forensic genetic laboratories should be prepared to cope with DVI situations, as this can involve large-scale DNA profile comparisons. Six forensic genetic laboratories from Switzerland participated in an exercise where supposedly a plane had crashed. The goal of the exercise was to monitor participants use of dedicated software with ground truth cases and to make them aware of the existence of particular situations that may occur in real cases. For assigning the value of the comparison of the DNA profiles, all participating laboratories used the DVI module of Familias v3.2.[1] In addition, one of the 6 laboratories used the Pedigree Searcher from CODIS v7.0. The data (AmpFlSTR® NGM SElect™ profiles) were generated to challenge the participating laboratories: cases with first, second degree biological parents, mutation events, as well as non-paternity cases were included. This study shows that the majority of the participants used the software in an appropriate way. However, a few misleading conclusions were detected for the most challenging situations. These errors belonged to one of the following categories: false pedigree, false association using the higher LR, misleading contextual information (false paternity) and not clustering family members. Specific recommendations are provided in order to reduce misuse of the software and the risk of misinterpretations by using all the relevant information.

1. Introduction

DNA analysis has been shown to be a very powerful tool to contribute to the identification of missing persons. In Switzerland, according to the Criminal procedure code,[2] forensic pathologists are in charge of this process in cases of unnatural death. Generally, they will combine information from different disciplines, such as forensic pathology, forensic odontology, forensic anthropology, fingerprints, DNA, in order to assign an identity to an unknown body.

The use of DNA profiling as a primary mean for victim identification in a mass disaster is well established and can be performed efficiently using the recommendations of the Interpol DVI (Disaster Victim Identification) Guide.[3] It is important to distinguish closed from open set forms of disasters. A closed set disaster is an event resulting in the death of individuals belonging to a fixed known identifiable group (e.g., passenger list), as it was the case for the Swiss Air Flight 111 that crashed in Nova Scotia, Canada in 1998.[4] A known airplane passengers' list was used for the identification of the two hundred and twenty-nine passengers and crew members. Such a configuration means that post-mortem (PM) data (i.e., samples from the deceased) can only come from the enlisted persons, for whom ante-mortem (i.e., samples from the alleged parent of the deceased or from her/himself when she/he was alive, AM)

* Corresponding author.
E-mail address: christian.gehrig@chuv.ch (C. Gehrig).

https://doi.org/10.1016/j.jflm.2021.102254
Received 26 May 2021; Received in revised form 15 September 2021; Accepted 17 September 2021
Available online 23 September 2021

will be generally available. An open set situation is quite different, as the event will have resulted in the deaths of an unknown number of individuals for whom, or part of whom, no prior records or descriptive data will be available. In such situations, it is often difficult to obtain information on the actual number of victims and their possible identity. The 9/11 terrorist attack on the World Trade Center in New York in 2001[5,6] and the 2004 tsunami disaster,[7] represent extreme examples of open set DVI situations. The latter disaster led the International Society for Forensic Genetics (ISFG) commission to publish recommendations regarding the role of forensic genetics for disaster victim identification (DVI).[8] These recommendations relate to the collection and storage of AM and PM samples, DNA extraction and genetic typing as well as data management, bio-statistical interpretation and the reporting of the results.

Switzerland has been so far spared from a large-scale DVI situation on its soil. Nevertheless, preparedness for a potential DVI situation involving a lot of DNA profile comparisons is essential in today's geopolitical situation. The GHEP-ISFG Working Group pointed out the importance of assisting DNA laboratories in gaining expertise in handling DVI or missing persons identification (MPI) through simulated DVI/MPI exercises.[9] The section of Forensic Genetics of the Swiss Society of Legal Medicine therefore decided, similar to the GHEP-ISFG Working Group, to organize a collaborative DVI exercise in order to monitor the participants conclusions and to make them aware of the existence of particular situations that may occur in real cases (mutation events, non-paternity and distant family relationship). Six of the seven laboratories from Switzerland, all accredited under ISO 17025 and authorized by the federal police, participated in an exercise involving a simulated plane crash. The plane crash scenario involved 87 bodies/remains as well as 97 biological relatives from 78 listed victims (i.e., a semi-closed set situation). Instructions for the exercise, as well as autosomal AmpFlSTR® NGM SElect™ (16 loci) alphanumerical profiles of the AM reference samples and PM samples were sent out to the participating laboratories.

The value of the comparison between AM and PM DNA profiles is generally quantified with a likelihood ratio (LR) given two mutually exclusive propositions and the case circumstances.[10] Briefly, a LR determines to which extent the DNA results support one proposition (e.g., the deceased is a child of Ms John) rather than an alternative proposition (e.g., the deceased and Ms John are unrelated). A likelihood ratio larger than 1 indicates that the DNA results are more likely if the first proposition is true rather than the alternative. A likelihood ratio smaller than 1 indicates a reverse situation. An LR close to 1, indicates that the DNA results cannot discriminate the two propositions, said otherwise the results are neutral. In order to assign one's probability of the proposition, one can use Bayes' theorem, which allows to update the probability of one proposition in the light of new information (here the DNA results). The LR represents the key component that allows to update our belief in the propositions or said otherwise to go from a "prior probability" to a "posterior probability". As an illustration, such an approach is used for paternity cases when using the Essen-Möller approach.[11] It considers that the 'prior' probability of the alleged father being the true father is equal to the 'prior' probability of an unknown, unrelated man being the true father.[1] In other words, the prior probability of paternity is assigned as 0.5. Using Bayes theorem with a prior probability of 0.5 and LRs of 10, 100, 500 or 1′000 lead to posterior probabilities of 0.909, 0.990, 0.998 and 0.999, respectively. This illustrates that one's posterior probability of the proposition (e.g., the deceased is a child of Ms John), depends not only on the LR but also on one's prior probability. The latter is based on non-DNA evidence, for instance - in DVI cases - the number of victims and/or other available forensic information (sex, age, clothes, localization of the body, etc.). Following the Swiss procedures within a DVI context, the DNA laboratory should assign a LR and the forensic pathologist, who has access to non-DNA evidence, should combine this LR with his/her prior probability and then decide whether the deceased is identified or not. This decision will depend on the posterior probability, but also on the consequences of identifying (or not) this person.[12]

Kinship cases and the assignment of LR can be difficult to handle without a dedicated software. Several programs exists for DVI and missing persons cases such as Familias (www.familias.no), Pedigree Searcher from CODIS v7.0, DNA-VIEW (http://dna-view.com/), Plass Data System (https://www.plassdata.com/products-services/software-products.html), Bonaparte (https://www.bonaparte-dvi.com/), M-FISys (https://www.genecodesforensics.com/software/).

This inter-laboratory exercise was the first of this kind organized in Switzerland. The main goal was to train for a DVI situation, to use a dedicated DVI software under supervision and to be aware of the common pitfalls. We did not intend it to be a proficiency test nor monitor whether the participants correctly 'identified' the persons. Indeed, we are of the opinion that DNA results on their own are not sufficient in order to take the decision to identify.

The participants were asked to contribute to the identification of the recovered bodies through family comparisons, including first degree relatives (biological parents, children, siblings), second degree relatives (grand-parents, half-siblings, uncle's), mutation events, partial profiles, re-association of body parts, as well as non-paternity.

2. Material and methods

2.1. Proposed simulated scenario

To challenge and train the scientists from the different Swiss forensic DNA laboratories, we provided the participants with instructions for the exercise, as well as autosomal AmpFlSTR® NGM SElect™ alphanumerical profiles of the AM reference and PM samples. The exercise consisted of the following semi-closed mock DVI scenario: a flight with 78 individuals on board had crashed shortly before landing in Geneva. At the crash site 87 bodies/remains were recovered, some of the bodies being dismembered. During the emergency landing, the plane had crashed into a house where a family of 4 (mother, father and their 2 children) were living, no reference samples were available for these victims. The participants were told that it was unknown if other causalities occurred. Hence they were unaware that 2 items, allegedly from 2 unknown persons present on the premises and killed by the impact of the plane, had been added. As AM data, DNA samples of 97 biological relatives (first and/or second degree) of the 78 missing persons were provided in order to contribute to identify the victims.

2.2. Generation of the DNA profiles

All profiles were generated using the program RStudio[13] and the toolbox 'DNAtools'[14] based on AmpFℓSTR® *NGM SElect*™ allelic proportions from Switzerland[15] with a co-ancestry coefficient of 0.01. The different pedigree trees were elaborated using this generated data. The participants received excel files with alphanumeric DNA data. A first file containing the DNA profiles of the recovered 87 bodies/remains (C profiles) was sent. In order to better simulate the results of an airplane crash, partial and full profiles were created.

Three direct and 94 familial reference DNA profiles were also generated in R. They were all listed in a second excel file that was intentionally partially filled out to facilitate the importation in the software used. The participants were asked to complete this file before launching the software. This was done with the help of a third file, the relationships table, summarizing the type of relationship expected between the victim and reference(s). In this table, different information was given: the family ID (F), the familial reference ID (R), the missing person's gender, the missing person's number, the relationship between

[1] This assumption is debatable, as shown by Biedermann et al. Equal prior probabilities: Can one do any better? Biedermann A., Taroni F., Garbolino P., 2007. Forensic Science International, 172 (2–3) pp. 85–93.

the victim and reference(s) and finally all useful comments about the pedigree (for example maternal grandmother, paternal uncle, etc). Supplementary material related to this article can be found, in the online version, at https://doi.org/10.1016/j.jflm.2021.102254.

2.3. Instructions to the participants

For the comparison of AM and PM DNA profiles, the 6 participating laboratories were asked to use the DVI module of Familias v3.2 software. The DVI module is divided into three steps, first the user adds the unidentified individuals/samples and their DNA profiles, second the reference families and the alleged pedigrees and last performs the DVI search. There are also several functions that may be carried out in each step, such as for example a *Blind search* (pairwise comparisons of different suggested relations) or applying different LR thresholds. One of the 6 laboratories additionally used the Pedigree Searcher from CODIS v7.0[16] in order to compare the results obtained with both software. Within this laboratory, the exercise was done independently using the two programs. We therefore considered the participation of 7 participants from 6 laboratories. CODIS v7.0 is a program containing enhanced missing persons and disaster victim identification tools, through the Pedigree searcher functions. It must be noted that this module is very efficient for ranking the persons according to the alleged relationship, but that the LRs calculated are not recommended for reporting purposes (Douglas Hares, personal communication). For the value of the results, operators must use the module PopStats and not the Pedigree searcher functions.

The participants were requested to use the Swiss population data,[15] a theta correction factor of 0.01, as well as the mutation model n° 1 *Equal probability (simple,* mutation rate of 0.001 for each genetic marker) in order to facilitate the comparison of the results.

For each re-association detected, the participants were asked to indicate their LR and whether they would report a "confirmed association", a "possible association" or "no association" as well as any useful comments (discrepancies, additional analysis needed, etc.) when reporting The participants were free to decide a "confirmed association" or a "possible association" (no LR thresholds were given) or just to give their LRs. The re-associations that remained undetected and those falsely reported were qualified as "missed association" and "false association", respectively.

Two participants had very limited experience with Familias, while the others used it routinely for example in complex kinship cases. Since this exercise, several new versions of Familias have been released with several improvements regarding the DVI process.

3. Results

3.1. Direct comparisons of remains (PM)

All participants first compared the 87 remains with each other, for example using "direct-match" in the *Blind search* function in Familias. This step allowed to help assign the number of victims. There were 3 re-association groups involving 6 items. All participants correctly re-associated the given profiles by direct comparison, assigning the number of victims as 84. Of the 84 victims, 2 bodies (C53, C78) did not have a familial reference because, unknown to the participants of the exercise, they were supposed to be present on the ground of the plane crash and thus additional causalities. A third body (C11) remained unidentified because of the non-paternity of the alleged biological father. The 81 remaining bodies were divided into 2 groups: group 1 consisting of pedigrees with multiple reference samples and/or first degree relatives and group 2 consisting of challenging cases with comparisons involving second degree relatives and/or mutations events and non-paternity.

3.2. Group 1: (comparisons with first degree relatives)

A total of 57 bodies with first degree references (parents, children and siblings) were proposed in this exercise. The two propositions used by the software for the interpretation were in each case either the body in question is related to the available reference sample(s) or the body is unrelated to the available reference sample(s). All the participants made the correct associations. The likelihood ratios obtained for a given relationship were all above 1000, meaning that the DNA results were at least 1000 times more likely if the alleged relationship between the body and the reference samples was true rather than if these persons were unrelated. Even though all samples were correctly associated, several errors were made in the establishment of the family pedigrees in Familias (Fig. 1) when using the automatic upload of AM information (profiles and relationships). These errors led to incorrect pedigrees, and therefore likelihood ratios with the propositions irrelevant to the case. The automatic upload is a tool to help the user, but, like any other software, pedigrees have to be carefully defined and verified. Version 3.2.8 of Familias incorporates an improvement to remind users of the pedigrees that have to be manually checked.

3.3. Group 2 (challenging comparisons)

Group 2 consisted of 24 bodies with essentially second degree relatives, mutations events and an unexpected non paternity case. In the vast majority of these 24 difficult cases, the comparison process was correctly handled by the participants. Most of the participants stated that they would have undertaken additional DNA testing (Y-STR's, mtDNA) to obtain more discriminating results.

3.3.1. Reporting comparisons considering possible mutational events

The exercise simulated a one-step maternal mutational event (C10), a one-step paternal mutational event (C47) and a two-step paternal mutational event (C48) between three bodies and their three parental references. Most of the participants accounted for possible mutation, and reported LRs of 26, respectively of 71 and 900 (usually, the LRs are rounded, but here we use the value obtained with the software in order to facilitate the comparisons of the participants answers). They correctly stated that the results supported the proposition that the deceased was related to the available reference sample(s) rather than unrelated. One participant (n°6) correctly considered a possible two-step paternal mutational event for C48 (with a LR of 900). S/he however failed to detect the other two associations C10 and C47 because the LR threshold was set at >100. As the two possible associations C10 and C47 showed LR values smaller than 100, namely of 26, respectively of 71 (Table 1), these possible associations were missed. The LRs assigned using CODIS Pedigree searcher function were always larger (less than factor of 5) than those assigned with Familias.

3.3.2. Reporting matches in case of siblings

In the case of victim C51, the so-called brother of the victim (in fact a half-sibling) was available as a familial reference. The expected LR considering that the persons were siblings or unrelated was 1. Two of the 7 participants considered that they could be half-siblings. With these propositions, an LR of 20 was obtained and victim C51 was correctly associated with his family. The other 5 participants did not report any association. In the case of victim C03, two brothers were available as familial reference. One of them was a true sibling, the other brother was unrelated to the 2 others. The participants were not told that one of the siblings was adopted and thus biologically unrelated to the other two individuals. Six of the 7 participants recognized this situation through the genetic relationship of the true sibling and the victim C03. All 6 participants assigned a LR in the order of 29′000 and issued a statement reporting this LR. Three participants indicated the association was possible and 3 as confirmed. Participant n°6 missed this association because the non-sibling was included in the pedigree (Table 1).

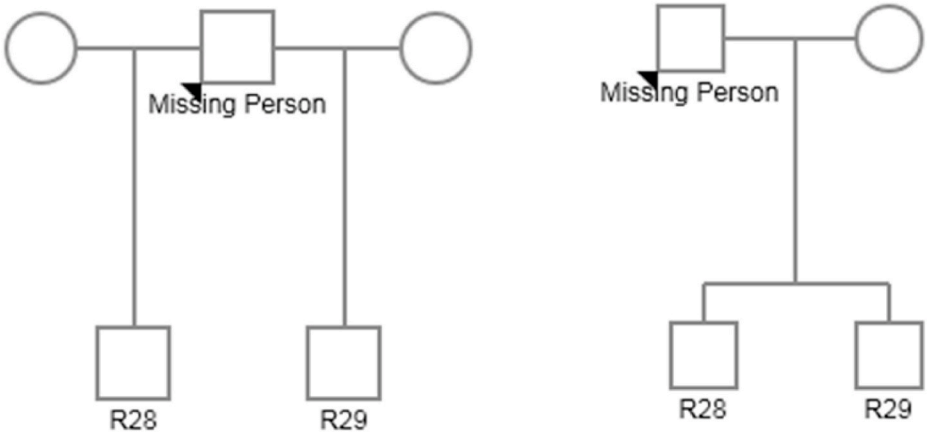

Fig. 1. Family 21, two sons (R28 and R29) are searching for their father. Familias 3.2 built, when using the automatic upload, the pedigree tree as half-brothers (left) instead of full-brothers (right) that was the correct situation.

Table 1
Likelihood ratios and conclusions reported by the participants for the challenging cases. The expected LR (Exp. LR) was assigned by the organizer. Participant n° 5 assigned its values with CODIS 7.0 whereas other participants used Familias 3.2.

Particularity	Relationship	Bodies	Exp. LR	Participants						
				n°1	n°2	n°3	n°4	n°5	n°6	n°7
1-step mutation	Maternal	C10	26	26^{P}	26^{P}	26^{P}	26^{P}	745^{P}	$-^{MA}$	26^{P}
1-step mutation	Paternal	C47	71	71^{P}	71^{P}	71^{P}	71^{P}	4990^{P}	$-^{MA}$	71^{P}
2-step mutation	Paternal	C48	900	900^{P}	900^{C}	900^{C}	900^{P}	$5.1E+04^{C}$	900^{C}	900^{C}
Distant relationships	Unexpected half-sibling	C51	1	20^{P}	20^{P}	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$
	2 brothers, only one true	C03	2.9E+04	$2.1E+04^{P}$	$2.9E+04^{C}$	$2.9E+04^{C}$	$2.9E+04^{P}$	$6.8E+05^{P}$	$-^{MA}$	$2.9E+04^{C}$
	Half-sibling	C64	81	81^{P}	$-^{NR}$	81^{P}	81^{P}	334^{P}	$-^{NR}$	81^{P}
		C62	115	115^{P}	115^{P}	115^{P}	115^{P}	505^{P}	115^{C}	115^{C}
		C39	29	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$	39^{P}	$-^{NR}$	$-^{NR}$
	Grandparent-child	C29	2	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$	21^{P}	110^{FA}	110^{FA}
		C04	14	$-^{NR}$	14^{P}	14^{P}	14^{P}	114^{P}	$-^{NR}$	$-^{NR}$
		C85	1	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$
		C60	72	72^{P}	$-^{NR}$	72^{P}	72^{P}	596^{P}	$-^{NR}$	72^{P}
	Avuncular	C12	7	$-^{NR}$	$-^{NR}$	$-^{NR}$	$-^{NR}$	40^{P}	$-^{NR}$	278^{FA}
		C27	734	734^{P}	734^{P}	734^{C}	734^{P}	6490^{P}	110^{C}	734^{P}
		C68	6	$-^{NR}$	$-^{NR}$	$-^{NR}$	6^{P}	27^{P}	$-^{NR}$	$-^{NR}$
Non Paternity	Maternal	C63	2E04	$2E04^{C}$	$2E04^{C}$	$2E04^{C}$	$2E04^{C}$	$2.3E05^{C}$	$-^{MA}$	$2E04^{C}$
Association through a living spouse	Husband	C77	2.90E+06	$1.6E+09^{C}$	$2.9E+06^{C}$	$1.6E+09^{C}$	$2.9E+06^{C}$	$2E+10^{C}$	$-^{MA}$	$2.9E+06^{C}$
	Son	C16	5.30E+04	$5.3E+04^{C}$	$5.3E+04^{C}$	$5.3E+04^{C}$	$5.3E+04^{C}$	$7.8E+05^{C}$	$5.3E+04^{C}$	$5.3E+04^{C}$
	Wife	C80	3.70E+05	$4.1E+07^{C}$	$3.7E+05^{C}$	$4.1E+07^{C}$	$3.7E+05^{C}$	$3.12E+08^{C}$	$-^{MA}$	$3.7E+05^{C}$
	Daughter	C35	7.00E+05	$7E+05^{C}$	$7E+05^{C}$	$7E+05^{C}$	$7E+05^{C}$	1.83E+07	$7E+05^{C}$	$7E+05^{C}$

^{MA}: missed association, ^{P}: possible association, ^{C}: confirmed association, ^{FA}: false association, ^{NR}: not reported.

3.3.3. Reporting the value of the comparisons in case of half-siblings

In 3 cases, a half-sibling was available as a familial reference sample (victims C64, C62 and C39). For victim C64, the expected LR was in the order of 80. Five of the 7 participants correctly associated victim C64 and indicated that the association was possible. The other 2 participants did not report any association. In the case of victim C62, the expected LR was 115. All 7 participants correctly associated victim C62, 5 indicated the association as possible and 2 as confirmed. In the case of victim C39, the expected LR was 30. Only one of the 7 participants correctly associated victim C39 and stated the association as possible. The other 6 participants did not report a possible association (Table 1).

3.3.4. Reporting the value of the comparisons when the reference is from grandparents

Four victims C29, C04, C85 and C60, had in each case, a single grandparent available as a familial reference sample. The first case C29, the expected LR was 2, participant n°5 assigned this LR for victim C29 and stated the association as possible. Four participants did not report any association. Participants n°6 and n°7 made a false association for the same body. The error consisted in falsely confirming the highest ranked body C41 as being associated with a LR of 110, as compared to the expected LR's of about 2 for the "true" family. This could have been avoided by sorting the results for the unidentified body C41. Indeed, this would have shown the high compatibility (LR in the order of 1 million), under a mother-child relationship, for the family F38. In the case of victim C04, 4 of the 7 participants assigned their LR as expected (LR = 14) and stated the association as possible. The other 3 participants did not report any association. In the case of victim C85, the expected LR was 1, none of the 7 participants reported a possible association. In the case of victim C60, the expected LR was in the order of 70, and 5 of the 7 participants correctly concluded that the results supported the propositions that victim C60 and reference R61 were related and stated the

association as possible. The other 2 participants did not report an association (Table 1).

3.3.5. Reporting the value of the comparisons in case of avuncular relations

Three victims C12, C27 and C68 had, in each case, a single avuncular parent available as a familial reference. The first case C12, the expected LR was 7 with family F11, participant n°5 correctly associated victim C12 with this family and reported the association as possible. Five participants (n°1 to n°4 and n°6) did not report an LR for that person with his true family. Participant n°7 made a false association between the family F11 and the body C43, rather than the victim C12. The error consisted in not considering comparisons alone and all the comparisons that had been done. This led to falsely confirming the highest ranked body as being a confirmed association with a LR of 278, as compared to the expected LR of about 7 for the "true" family. This could have been avoided by sorting the results for unidentified body C43. Indeed, this would have shown a high compatibility (LR in the order of 1 million) with body C84 of the family F71 (family of four living in a house destroyed by the plane crash). In the case of victim C27, the expected LR was in the order of 700, all 7 participants correctly reported an LR in the order of 700 (respectively of 6500 for the participant using CODIS (n°5)) supporting that victim C27 and family F27 were related rather than not. Although participant n°6 correctly reported that the results supported that victim C27 was related to family F27, they assigned an "incorrect" LR of 110. The error consisted in forgetting to add a grandfather in the pedigree. In the case of victim C68, the expected LR was 6, 2 of the 7 participants reported this LR for victim C68. The other 5 participants did not report any association, nor LR for the ground truth proposition (Table 1).

3.3.6. Reporting results in case of a non-paternity

For victim C63, the mother and the father were available as familial references. The mother was truly the biological mother of C63, but the presumed father was not the biological father of C63. The expected LR considering both parents is 0 whereas the LR considering only the mother is 20′000 (Table 1). Six of the 7 participants recognized this non-paternity through the genetic relationship of the mother and victim C63. Thus, they issued a statement reporting results supported an association. Participant n°6 did not report this association because of the non-paternity (Table 1).

3.3.7. Reporting results for a spouse/parent

Two pairs of victims C77/C16 and C80/C35, in each case a parent and a child, were to be suggested as possible relatives through a familial reference consisting of a spouse/parent (Table 1). In the case of victims C77 and C16, the wife/mother (R97) of 2 victims was available as a familial reference. The expected LR's were 2.9E+06 (if C77 is or not the father of the child C16), respectively 5.3E+04 (R97 is or not the mother of C16), 6 of the 7 participants had these LR's or larger and correctly concluded that the results supported that the 2 victims were from the same family. Participants n°1 and n°3 had higher LR's because after assuming that there was an association between the child with the mother, this information was used to create a new pedigree (trio). In the case of victims C80 and C35, the husband/father (R51) of 2 victims was available as a familial reference. The expected LR's were 3.7E+05 (if C80 is or not the mother of the child C35), respectively 7E+05 (if R51 is or not the father of C35), 6 of the 7 participants reported these LRs or larger. Participants n°1 and n°3 had larger LRs because once the child was assumed to have been identified, the information was used to create a new pedigree (trio). Participant n°6 did not report the compatibilities (nor their LR) for the 2 spouses because they did not use the *Blind search* function between the unidentified bodies.

3.4. Reporting results in case of a family with no reference material

A family of four (F71), 2 parents, their daughter and their son were to be identified by the forensic pathologist, but no reference samples were available (Table 2). Four participants (n° 1, 2, 4 and 5) correctly suggested that these 4 victims could be F71. Participants n° 3 and n°7 correctly associated the 4 victims to the same family but did not distinguish the 2 female victims as the mother or the daughter, respectively the father or the son. Participant n°6 missed suggesting this lead, because they did not use the *Blind search* function between these unidentified bodies.

4. Discussion

In this simulated DVI exercise, we chose different degrees of relationship in order to test and train the participants in their use of the software (Familias v3.2 or Codis Pedigree searcher). Related victims, mutations events and partial profiles were added so that the exercise would present a variety of situations that could be encountered in casework. Despite the complexity of the exercise, most laboratories used the software in the intended way and reported LRs that supported the ground truth proposition. This is especially true in situations where reference samples from first-degree relatives were available (57/84 cases). For the three bodies where no relevant familial reference was available, no information was obviously provided by the DNA results. The situation was more challenging for the remaining 24 bodies, because of the introduction of mutation events or because the comparison had to be done with more distant relatives. In these more difficult situations, a body was suggested as a possible association with the wrong family in 3 cases. Associations were missed in 6 cases. Interestingly, these errors only concerned the two participants who had less experience with the software. The false associations were explained because two participants had, for the two cases (C12 and C29), reported the association with the largest LR whereas the ground truth person presented a smaller LR. Sorting the results by unidentified bodies helps to find out if a body has been assigned LRs larger than 1 with more than one family.

When considering parentage cases, it is always a safe course of action to consider the possibility that the true genetic relationships can be different from the reported genetic relationship. For family F09, 2 references were available: mother R11 and a "non-biological father" R12. If one does not consider the possibility that there might be an error or that the father might not be the biological father, no association is suggested: as indeed R12 is not the biological father, whereas the mother is. When comparing pair of DNA profiles, for instance with the *Quick Scan* function of Familias, a parent-child search, the LR obtained supports the proposition that R11 is the mother.

On the other hand, making use of all the information in a pedigree, is advised to increase sensitivity and specificity. For families where several members are missing, it is advised (if a possible association is observed) to use this information to complete the pedigree used for other missing members. Those using this strategy[10] were able to report findings that supported the ground truth proposition. For example, for family F29, there are 3 full siblings (C22, C30 and C37) missing and the reference person is the father R40. When considering only the child and father in the pedigree and a LR threshold of 100, 2 bodies are suggested as possible children: a man C22 and a woman C30. These 2 associations are

Table 2
Results of the *Blind search* of the family of four victims (two parents and their daughter and son without reference material).

Relationship	Body	Participants		
		n°1, 2, 4, 5	n°3, 7	n°6
Mother	C04	C04[C]	C04 or C33	-[MA]
Father	C84	C84[C]	C43 or C84	-[MA]
Son	C43	C43[C]	C43 or C84	-[MA]
Daughter	C33	C33[C]	C04 or C33	-[MA]

[C]: confirmed association, [MA]: missed association.

also suggested by the sibling's *Blind search* results. For C37, the third sibling, adventitiously there are only a few alleles that are shared, hence the LR was smaller than the threshold of 100. By grouping family members, for instance with the *Move* function in Familias, one can associate bodies C22 and C30 to the family label F29P3. With this updated configuration, the third child (C37) has a large LR supporting the ground truth proposition.

Participants have used different thresholds to distinguish situations where the associations were not reported, possible or confirmed, which reflects different practices. One could argue that all LRs should be reported, but in a DVI situation this is not feasible. Most reported a possible association when their LR was between 1 and 100. Therefore, some associations were reported by some laboratories and not by others. 'Confirmed association' were used when the LR was higher than 100, 500 or 1000, depending on the laboratory and the relationship. Consequently, different national laboratories may transmit different conclusions despite the value of the DNA evidence being the same. This shows the importance of being transparent regarding the value of the results (i.e., to report one's LR value whatever the number).

All laboratories used the LR to assign the value of the DNA results. No laboratory reported a posterior probability. Nevertheless, we asked the participants what would have been their assignment for their prior probability in such a situation. Most would have used 1/84, 84 corresponding to the number of different unidentified persons. One participant would have used a prior probability of 1/82 and the last two participants a prior probability of 1/87. This last prior probability seems more difficult to justify since it was based on the amount of different remains (87) and not on the different number of unidentified persons (N = 84). Using different prior probabilities may further induce variation between the results from different laboratories. In any case, the impact of these different prior probabilities (1/82, 1/84 or 1/87) is low. Furthermore, the DNA laboratory has generally little if any information concerning the non-DNA evidence. Nor is it their duty. As shown by Bayes theorem, we believe that LRs are the responsibility of the laboratory and prior/posterior probabilities, the responsibility of the identification board.

It must be noted that Familias computes likelihoods, likelihood ratios and posterior probabilities. In version 3.2, in a closed set, posterior probabilities do not correspond to what would be obtained if the user externally calculated posterior odds by multiplying prior odds (assigned as 1/N) by Familias computed LR and then converting these posterior odds into a posterior probability (see Appendix). Familias 3.2 computed posterior probabilities (in a closed set) take into account the information provided by the other comparisons. This can help identify the problematic situations such as the ones described. In newer versions of Familias, prompted by our discussions with the authors of this program, one can choose different ways to obtain posterior probabilities, depending on which propositions are meaningful in the case. This can help identify the problematic situations such as the ones mentioned for C12 and C29.

Regarding the communication of the results, we would like to comment the Interpol DVI forms that are provided for the transmission of the DNA results. Indeed, there is no possibility that would allow to report the LR, and the only boxes that can be ticked are: not relevant/no data available or insufficient data/identification possible/probable or established (Fig. 2).

This shows the misunderstanding of the role of the DNA examiner and a change would be most welcomed to adhere to professional guidelines [ENFSI, ISFG].

To prevent the incorrect or missed associations highlighted above when handling a DVI case with a dedicated software, we outline a possible procedure. The recommendations are described below with a particular emphasis on Familias and CODIS (Pedigree searcher).

4.1. Merging the results of remains assumed to have come from the same individual

After the import of the DNA profiles of the remains, the first step is to determine whether some could be from the same person. For example, it was stated in our scenario that 78 people were onboard, when the plane crashed into a house where a family of four lived. A minimum of 82 different DNA profiles were expected. A direct *Blind search* match with Familias or an identity search with CODIS should be performed to confirm (or not) this assumption. According to the search conducted, three pairs of remains were associated. After merging the three pairs, 84 different DNA profiles were available. The previous assumption accounted for 82 individuals. Therefore, there were at least 2 unannounced bodies (C53 and C78) without relative's reference samples. Here, we do not consider the possibility of there being twins. It would be important in the DVI DNA report to disclose the assumptions that have been made, such as merging samples sharing similar profiles.

4.2. Controlling the pedigree established by the software

After importing the DNA profiles of the remains, it is necessary to import the DNA profiles of the references and check that the pedigrees established by Familias or CODIS correspond to the case situation.

In this study, use of partially adequate pedigrees had a very low impact on the value of the results. The two participants who did not check the pedigrees obtained LRs that were of the same order of magnitude as the ones expected. However, this ought to be avoided.

4.3. Identifying potential difficulties

Once pedigrees have been controlled or manually created, potential difficulties can be anticipated. First, one should assess the number of families present among the missing. By observing the file of relationships, it is possible to determine how many families have several missing members. Using the *Blind search* function in Familias (or similar) and parent-child, as well as sibling's relationships, two lists of potential compatibilities among the bodies can be generated. This enables to use all the available information and optimizes sensitivity and specificity. To investigate parent-child relations of bodies, it is possible to use the CODIS Identity Search with low stringency (in contrast to the CODIS Pedigree Search).

Knowing how complex the situation is and how many different types of family relationships there are can also be useful. There were 58 parent-child cases (3 involved a mutation and in 4 cases alleged fathers were not the biological fathers). For other relationships, there were 28 full siblings, 3 half-siblings, 4 grandparents-grandchildren, 3 uncle-nephew and 1 wife and 1 husband missing. It is known that relationships such as half-siblings[2] may provide limited information. If there is only one person whose reference is available, then it is fairly straightforward for a given kit to decide if further references are needed. But, if there are more than one, it becomes more difficult. Familias can use the genotypes of the references available for each family and generate the expected LRs by computer simulations. This allows the user to pre-assess the case and generate the expected LR (e.g., 100) for each family if the deceased is the family member. These simulations showed that 17 families were not ideal in terms of degree of relatedness. As expected, half-siblings, grandparents and uncle relationships are generally characterized with LRs smaller than 100. Once this information is known, one can ask for other references. With the simulations, the user is also made aware of the fact that setting for example an LR threshold at 100 can lead to misleading conclusions for these 17 families. CODIS software does not allow to pre-assess the case. Note that the present study focused

[2] The probability of having one or no allele identical by descent is the same for half-siblings, grandparents-grandchildren or uncle-nephew.

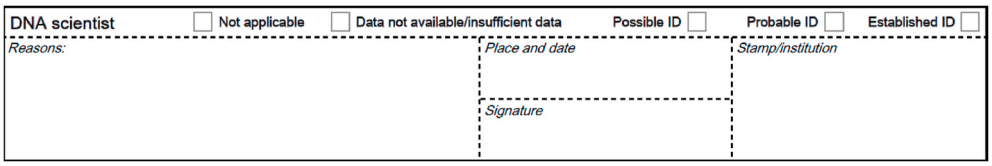

Fig. 2. Page 1 of the Interpol identification report (version 2018).

on 16-STR loci autosomal DNA profiles. Presently, the laboratories in Switzerland are all using 23-STR autosomal loci (Powerplex® Fusion 6C amplification kit) for cases where the issue is the identification of an unknown dead body. This increases the sensibility and specificity of the DNA analysis.

4.4. Association of missing children with their potential parents

To help associate missing children with their parent(s) there are three strategies. The first one is to use the search function and set the same LR threshold for all relationships. The second is to use the selected search, which allows working only with one family to find the possible missing person. This allows detecting mutations more easily. A further possibility is to perform a *Quick Scan* in Familias. Since it is a pairwise comparison, the *Quick Scan* allows bypassing cases with differences between the actual and the reported genetic relationship (e.g., non-paternity in a trio). By comparing results with the relationship file, the proposed associations can be sorted and studied. In order to verify the compatibilities and help associate them with the correct family label we also need to check the gender of the remains if the software does not take this information into consideration. Once a possible association is detected, three choices are available. The type of strategy used ought to be mentioned in the report (in the methodology for example).

a. The user can remove the body associated from the list of bodies or move it into 'its family' by associating it with its family label (for example F1P1). This operation is irreversible. This operation is a good option if there is only one missing member in the family. This action will reduce the list of other future potential associations.
b. The user can leave the body in the list without associating it to a family, to compare it with other potential candidates. When the LR obtained is under a predetermined threshold it is safer to let the body in the list for further comparisons.
c. If there are several missing members from the same family, it can be advantageous to use the *Move* function in Familias to place the body into the family label of a relative. This allows to use all the pedigree and by adding relevant information we achieve better sensitivity and specificity. For example, this strategy allows the association of a body with the wife of F37 and with the husband of F70 through their children.

5. Conclusion

In conclusion, our study highlights that DVI situations can quickly become complex and the use of dedicated software to manage DNA data comparison is essential. Nevertheless, as pointed out by Vullo and colleagues[9] the use of a powerful software does not ensure correct results. Proper training, validation, proficiency testing or collaborative exercises, as well as regular use of the software are essential to handle complex DVI situations. It is therefore important that exercises such as the one presented here are regularly organized in order to check that the results obtained are the ones expected. These exercises should simulate real situations and ideally include the different specialists involved in DVI cases.

A special attention should be given to the following points to reduce the risk of misleading results:

- Disclose the assumptions used.
- Verify the pedigrees generated by the software.
- Merge remains that are assumed to have come from the same individual.
- Consider the whole pedigree after having checked them individually.
- Be aware that the true and the reported genetic relationship might be different. One shall use all the information available, but also ensure that this is relevant information. To take into account the possibility that the biological relationship is different from the civilian one, we recommend doing a pairwise comparison between each body with the different familial references: otherwise misleading conclusions could be given.
- Be cautious with low LRs (<500) and remember that whatever the order of magnitude of the LR, the highest value is not always the true genetic association (and this is also why we do not identify with DNA only).
- Mutation(s) may explain situations where a low LR is obtained when comparing potential first-degree relatives. Mutations are expected to occur only at a few loci.
- Be cautious in case of open or semi open DVI situations, because several of the DVI victims will be lacking familial references.
- When several laboratories are involved in a case, provide instructions regarding the parameters to use in order to homogenize the transmission of the results (Population data, F_{ST} value, LR thresholds).
- If needed, perform extra DNA analyses. Lineage markers located on the Y-chromosome or on the mitochondrial DNA may provide useful information on the paternal and maternal relationships, respectively.
- When available, samples from first-degree relatives should be preferred. Personal references can provide useful information if their origin is clearly established, for instance a biopsy, a blood sample from an alcohol test or a Guthrie card.[17]
- Forensic DNA scientists should not conclude that an identity is confirmed.

Finally, this study highlights that the person who is in charge of the identification of an unknown body should be aware that DNA results alone are not sufficient for identification and can be misleading in rare situations. This is particularly the case when the strength of the DNA evidence is low. In such cases, it is recommended to collect more information (see above). Furthermore, combining the DNA results with the available non-genetic information (e.g., gender, fingerprints, age and medical records) should increase the probability that the person who is in charge of the identification makes the right decision.

Declaration of competing interest

None.

Acknowledgements

We are extremely grateful to all participants who agreed to take part in this exercise and thank the reviewers for their useful comments. We also wish to thank Daniel Kling for our useful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jflm.2021.102254.

References

1. Kling D, Tillmar AO, Egeland T. Familias 3-Extensions and new functionality. *Forensic Sci. Int. Genet.* 2014;13:121–127.
2. Art, 253 https://www.fedlex.admin.ch/eli/cc/2010/267/en.
3. https://www.interpol.int/content/download/589/file/18Y1344%20E%20DVI_Guide.pdf.
4. Leclair B, Frégeau C, Bowen K, Fourney R. Enhanced Kinship Analysis and STRbased DNA typing for human identification in mass fatality incidents: the Swissair Flight 111 disaster. *J Forensic Sci.* 2004;49(5):939–953.
5. Biesecker L, Bailey-Wilson J, Ballantyne J, et al. DNA Identifications after the 9/11 world trade center attack. *Science.* 2005;310:1122–1123.
6. reportLessons Learned from 9/11: DNA Identification. The report is available at: http://www.ncjrs.gov/pdffiles1/nij/214781.pdf.
7. Lau G, Tan WF, Tan PH. After the Indian Ocean tsunami: Singapore's contribution to the international disaster victim identification effort in Thailand. *Ann Acad Med Singapore.* 2005;34:341–351.
8. Prinz M, Carracedo A, Mayr WR, et al. DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI). *Forensic Sci. Int. Genet.* 2007;1:3–12.
9. Vullo CM, Romero M, Catelli L, et al. GHEP-ISFG collaborative simulated exercise for DVI/MPI: lessons learned about large-scale profile database comparisons. *Forensic Sci. Int. Genet.* 2016;21:45–53.
10. Aitken CG, Taroni F, Bozza S, eds. *Statistics and the Evaluation of Evidence for Forensic Scientists*. third ed. John Wiley and Sons Ltd; 2021.
11. Essen-Möller E. Die Beweiskraft der Ähnlichkeit im Vaterschaftsnachweis - theoretische Grundlagen. *Mi.tteilungen de Anthropologischen Gesellschaft.* 1938;68:9–53.
12. Biedermann A, Vuille J. *Understanding the Logic of Forensic Identification Decisions (Without Numbers)*. Sui-generis; 2018:397–413.
13. RStudio Team. *RStudio*. Boston: Integrated Development for R. RStudio, Inc.; 2016. http://www.rstudio.com/.
14. Tvedebrink T, Curran J. *DNAtools : Statistical Functions for Analysing Forensic DNA Databases*; 2011. R Package Version 0.1-2 http://www.cran.rproject.org/web/packages/DNAtools.
15. Gehrig C, Balitzki B, Kratzer A, Cossu C, Malik N, Castella V. Allelic proportions of 16 STR loci—including the new European Standard Set (ESS) loci—in a Swiss population sample. *Int J Leg Med.* 2014;128:461–465.
16. CODIS 7.0. https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet.
17. Hartman D, Benton L, Morenos L, Beyer J, Spiden M, Stock A. The importance of Guthrie cards and other medical samples for the direct matching of disaster victims using DNA profiling. *Forensic Sci Int.* 2011;205:59–63.