Forensic Science International: Genetics 2021; 54:102519 pp. 1-8
(DOI: 10.1016/j.fsigen.2021.102519)

Making decisions in missing person identification cases with low statistical power

Franco L. Marsico[1,*], Magnus D. Vigeland[2], Thore Egeland[3] and Mariana Herrera Piñero[4]

[1]Banco Nacional de Datos Genéticos, Avda. Córdoba 831, C1054AAH-CABA, Argentina. Email: franco.lmarsico@gmail.com

[2]Department of Medical Genetics, University of Oslo, PB 4956 Nydalen, 0424 Oslo, Norway. Email: magnusdv@medisin.uio.no

[3]Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Chr. M Falsens vei 1, 1433 Aas, Norway. Email: thore.egeland@nmbu.no

[4]Banco Nacional de Datos Genéticos, Avda. Córdoba 831, C1054AAH-CABA, Argentina. Email: mherrera@mincyt.gob.ar

(*) Corresponding author.

Highlights

- General strategy for dealing with low statistical power in missing person and disaster victim identification cases.
- **Method for selecting a likelihood ratio (LR) threshold for identifications through DNA-database search.**
- Error rates estimated by conditional simulation.
- **Examples based on unsolved cases of '*Missing Grandchildren of Argentina*'.**
- Freely available software.

Making decisions in missing person identification cases with low statistical power

Abstract

The present work proposes a general strategy for dealing with missing person identification cases through DNA-database search. Our main example is the identification of abducted children in the last civic-dictatorship of Argentina, known as the "Missing Grandchildren of Argentina". Particularly we focus on those pedigrees where few, or only distant relatives of the missing person are available, resulting in low statistical power. For such complex cases we provide a statistical method for selecting a likelihood ratio (LR) threshold for each pedigree based on error rates. Furthermore, we provide an open-source user friendly software for computing LR thresholds and error rates. The strategy described in the paper could be applied to other large-scale cases of DNA-based identification hampered by low statistical power.

Introduction

The use of DNA databases for missing person identification (MPI) and disaster victim identification (DVI) has gained relevance in the last decades [1,2,3]. Despite the incorporation of more genetic markers and new software programs for kinship testing, several complicating factors remain challenging in such applica- tions, mainly pedigrees with low statistical power, complex pedigree structures and inconsistencies due to mutations.

Loosely speaking, the statistical power refers to the probability of reaching a conclusion when we are testing a specific hypothesis with the available data. In our context, the data is the genetic information collected from: (i) persons of interest (POIs) that are people who could be the missing person (MP) and (ii) reference individuals, i.e., relatives of the missing persons. Typically, each POI is matched

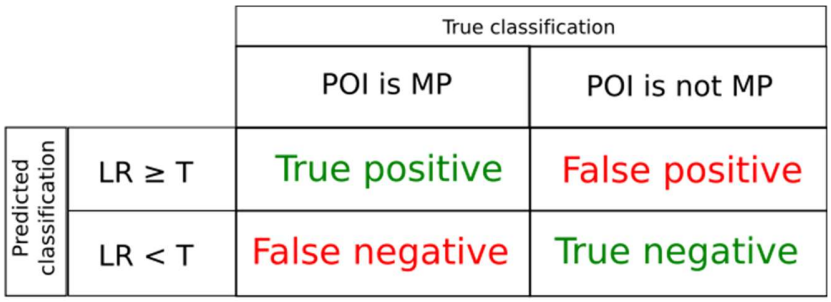

Figure 1: Confusion matrix showing potential outcomes of LR based classification.

against the reference families using a statistical kinship test based on the likelihood ratio (LR). A potential match is declared if LR exceeds a specified threshold *T*.

Two measures of statistical power are particularly useful in large-scale cases such as the "Missing Grand- children of Argentina" [1,4]: The inclusion power (IP) is the probability that LR $\geq T$ if POI is MP, while the exclusion power (EP) is the probability that a random unrelated person can be excluded due to genetic inconsistencies with the relatives of the missing person (mutation model is not used when estimating EP). Low statistical power increases the risk of not reaching the correct decision [5] (Figure 1). Potential matches, indicated by LR $\geq T$, are analysed further, with more data (mtDNA, Y-STRs or X-chromosomal markers), to reach a conclusion about the case. This confirmation step has the specific aim of avoiding false positives. In contrast, false negatives (LR $< T$, when POI is MP) are silent errors which are often hard to catch.

Choosing an optimal LR threshold T is a difficult problem in many forensic applications. In practice, DNA laboratories set an $ad - hoc\ T$ based on the type of cases that they are working on (MPI, DVI, paternity testing, etc). In MPI cases, high thresholds are generally chosen in order to avoid false positives. For example T = 10,000 was proposed by Kling et al. [1] where a set of unsolved cases from BNDG were exemplified. After performing simulations with a group of reference pedigrees from BNDG, authors did not get false positives, using T = 10,000, for all analyzed cases. Importantly, for

pedigrees with low statistical power, the threshold of 10,000 is not likely to be met; in other words such pedigrees have a high false negative rate (FNR). Selecting a lower T decreases FNR at the cost of increasing the false positive rate (FPR). On the other hand, a high FPR due to a low T implies an increased cost for a forensic lab, since extended analyses are needed to check each potential identification.

Previous works have explored some strategies for selecting T based on LR distributions [6,7,8]. These approaches analyzed different types of relationships (uncle-nephew, grandparent-grandchildren, etc) in order to determine LR thresholds for declaring an identification. As mentioned by Slooten et al. [5] a distinction should be done between notions of evidence and decision making in these cases. We do not provide a LR threshold for reaching a conclusion in the identification process, when comparing H1: POI is MP, and H2: POI is unrelated to MP. We propose a case-specific LR threshold, named LR decision threshold (DT), for reference families with low statistical power to identify an optimal subset of POIs. Similarly, Meester et al. suggested using a case specific LR threshold based on error rates for selecting this subset [8]. Here, we propose a statistical method for selecting a case-specific DT. Furthermore, we exemplify its application on the Banco Nacional de Datos Genéticos (BNDG) cases.

The "Missing Grandchildren of Argentina" is a pioneer case in missing person identification, and BNDG, created in 1987, is one of the first of its kind [9,10,11]. It has two genetic databases: one for the reference pedigrees composed by the genetic profiles of the relatives of the missing persons and the other for people who question their biological identity and were born during the military-civic dictatorship, namely, persons of interest (POIs). DNA database search implies performing kinship analyses between all POIs against all the pedigrees present in BNDG. A majority of BNDG reference pedigrees lack first-degree relatives of the missing person, because biological parents of the kidnapped children were murdered and their bodies still remain missing. This implies that the DNA-based identification of the kidnapped children must be done with distant relatives, like grandparents, uncles, great-grandparents, and others [12]. These characteristics of the "Missing Grandchildren of Argentina" are not unique. There are several cases where first degree relatives of the missing persons are

unavailable for genotyping [3]. In such complex situations the more distant relatives play a key role in the identification process. Likelihood ratio Decision Threshold calculation and estimation of error rates can be obtained using Forensic Genetic Simulation Analysis (FOGSA), developed with the aim of helping decision makers in database search analysis.

Materials and methods

Database. BNDG currently holds genetic data from more than 10,000 POIs. In addition, each month around 100 new POIs are incorporated into the database. The reference database contains around 300 pedigrees composed by relatives of the missing grandchildren. This study covers 274 of the reference families. In the last general evaluation of the BNDG [1], summary statistics of the analyzed reference families were calculated providing an overview of the number of typed relatives and their relationship to MP. Since then, the number of genetic markers has increased from 15 to 23 for all reference samples in the database. For special cases, when more genetic data is required, some pedigrees are genotyped with more autosomal STRs markers (increased to at least 30), X-chromosomal STRs, Y-STRs and mtDNA. **Kinship testing**. The identification cases considered in this paper involve families with a single MP. For a given POI we compare the following hypotheses: H1: POI is MP and H2: POI is unrelated to MP. After genotyping POI, a statistical evaluation of the above hypotheses is done using the likelihood ratio, defined by

$$\mathrm{LR} = \frac{P(data \mid H1, \theta)}{P(data \mid H2, \theta)} \qquad (1)$$

Here data means the genotype information, and ϑ contains marker properties and other fixed parameters. In our context each hypothesis H is equivalent to specifying a (not necessarily connected) pedigree, and we may regard $P(data|H)$ as a pedigree likelihood.

Statistical Power Analysis. Statistical power for kinship testing was described by Egeland et al. [13]. Kling et al. [1] presented power calculations based on conditional simulations, currently implemented in the R library forrel (https://CRAN.R-project.org/package=forrel). Unless otherwise stated, we assume

that the reference data is consistent, meaning that the observed genotypes have nonzero probability in the reference pedigree (but note that this may allow mutations if appropriately modelled) [4]. For the simulations in this paper we applied the uniform mutation model, with a mutation rate of 0.002 [14].

Based on the previously mentioned work by Kling et al. [1], we defined the inclusion power as the probability that LR exceeds T_0 = 10,000, given that POI is MP:

$$\mathrm{IP}_{10,000} = P\,(\mathrm{LR} \geq 10{,}000 \mid H1) \tag{2}$$

We estimated $\mathrm{IP}_{10,000}$ for each pedigree by simulating the genotypes of 10,000 MPs conditional on the reference data.

The exclusion power (EP) is the probability of LR vanishing, given that POI is not MP:

$$\mathrm{EP} = P\,(\mathrm{LR} = 0 \mid H2) \tag{3}$$

Reference pedigrees with the same structure (for example, one uncle is typed) may have different statistical power, depending on the observed genotypes. The power is typically better for pedigrees composed by members with rare multilocus genotypes, compared to more common genotypes. To account for such differences, all simulations were performed conditionally on the known reference data.

We may consider the false negative rate as a function of the threshold T,

$$\mathrm{FNR} = P\,(\mathrm{LR} < T \mid H1) \tag{4}$$

and similarly for the false positive rate:

$$\mathrm{FPR} = P\,(\mathrm{LR} \geq T \mid H2) \tag{5}$$

For each pedigree, with conditional simulations, we estimated FNR and FPR for each integer T between 1 and 10,000. For cases where low rates were obtained, importance sampling was used [15].

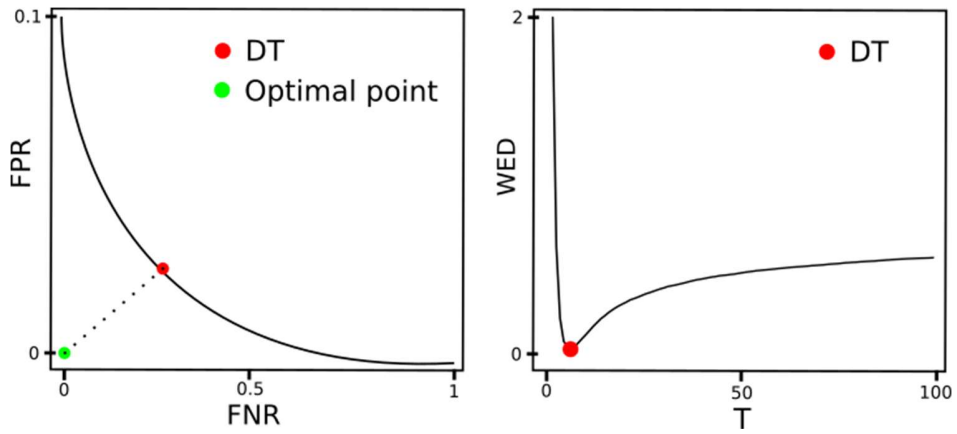

Figure 2: Left. False negative rate (FNR) and false positive rate (FPR) when the LR threshold T varies between 1 and 10,000. The red dot minimizes the weighted euclidean distance (WED) to the origin. Right. WED as a function of T.

LR decision threshold estimation. DT is the value of T that minimizes the following weighted euclidean distance (WED) to the theoretical optimum (where FPR = FNR = 0) (see Figure 2):

$$\mathrm{WED} = \sqrt{(W_1\mathrm{FPR})^2 + (W_2\mathrm{FNR})^2} \tag{6}$$

The choice of weights W_1 and W_2 reflects the relative importance of false positives and negatives, and depends on the database size and the capacity of the forensic lab. A ratio of ten to one, corresponding to $W_1 = 10$ and $W_2 = 1$, between FPR and FNR allows obtaining DTs with a manageable number of expected false positives.

We note that a similar approach, the point closest to the optimal, is commonly used in the estimation of optimal cutoff in ROC curves [16]. Here, the weights represent the relative importance of the errors and depend on laboratory resources. The proposed weights in this paper were selected for the particular cases from BNDG but may be modified for decision makers in their cases. This is addressed in more detail in the discussion section.

Database search strategy. A series of steps, illustrated in Figure 3, were followed for each pedigree:

(i) Statistical power evaluation: The first step is the evaluation of the statistical power of the reference

pedigree. In this way we are able to identify well-powered (EP ≥ 0.9; $IP_{10,000}$ ≥ 0.9) and underpowered pedigrees (either EP or $IP_{10,000}$ below 0.9). For pedigrees with good power, we go to the step (iv). For those underpowered we follow the steps detailed below.

(ii) Include more relatives: If there are available relatives, the impact of their incorporation to the pedigree is tested and the best option is genotyped (for more details see Vigeland et al. [4]). For pedigrees that now reach good power, we go to step (iv); otherwise we proceed with the next step.

(iii) Likelihood ratio decision threshold calculation: A decision threshold DT is computed as explained above.

(iv) DNA database search: Cases with LR ≥ 10,000 (well-powered pedigrees) or LR ≥ DT (for underpowered pedigrees) are declared as potential identifications.

(v) Confirmation: All potential matches are analyzed in detail, possibly with more data incorporated, to reach a reliable conclusion. If all collected evidence is consistent, the match is confirmed and declared as a missing person identification.

Implementation. The software Familias was used to perform the LR calculations [17,18]. Plotting and power analyses were done in R with the forrel package (https://cran.r-project.org/web/packages/forrel/index.html). Importance sampling, FNR, FPR and DT calculations were done using the mentioned FOGSA (https://github.com/MarsicoFL/FOGSA). A web version is available (https://francomarsico.shinyapps.io/fogsa/) as well tutorial and documentation: https://marsicofl.github.io/FOGSA/. This software works with the output data from simulations performed with programs such as Familias and forrel.

Results

Power analysis for DNA-based identification. Figure 4 shows a power plot with EP and $IP_{10,000}$, for all analyzed BNDG reference families. With 23 markers, more than 70% of the pedigrees present in the BNDG have good statistical power (EP and $IP_{10,000}$ both ≥ 0.90).

Supplementary Table 1 shows that the median $IP_{10,000}$ (considering all the pedigrees in the reference

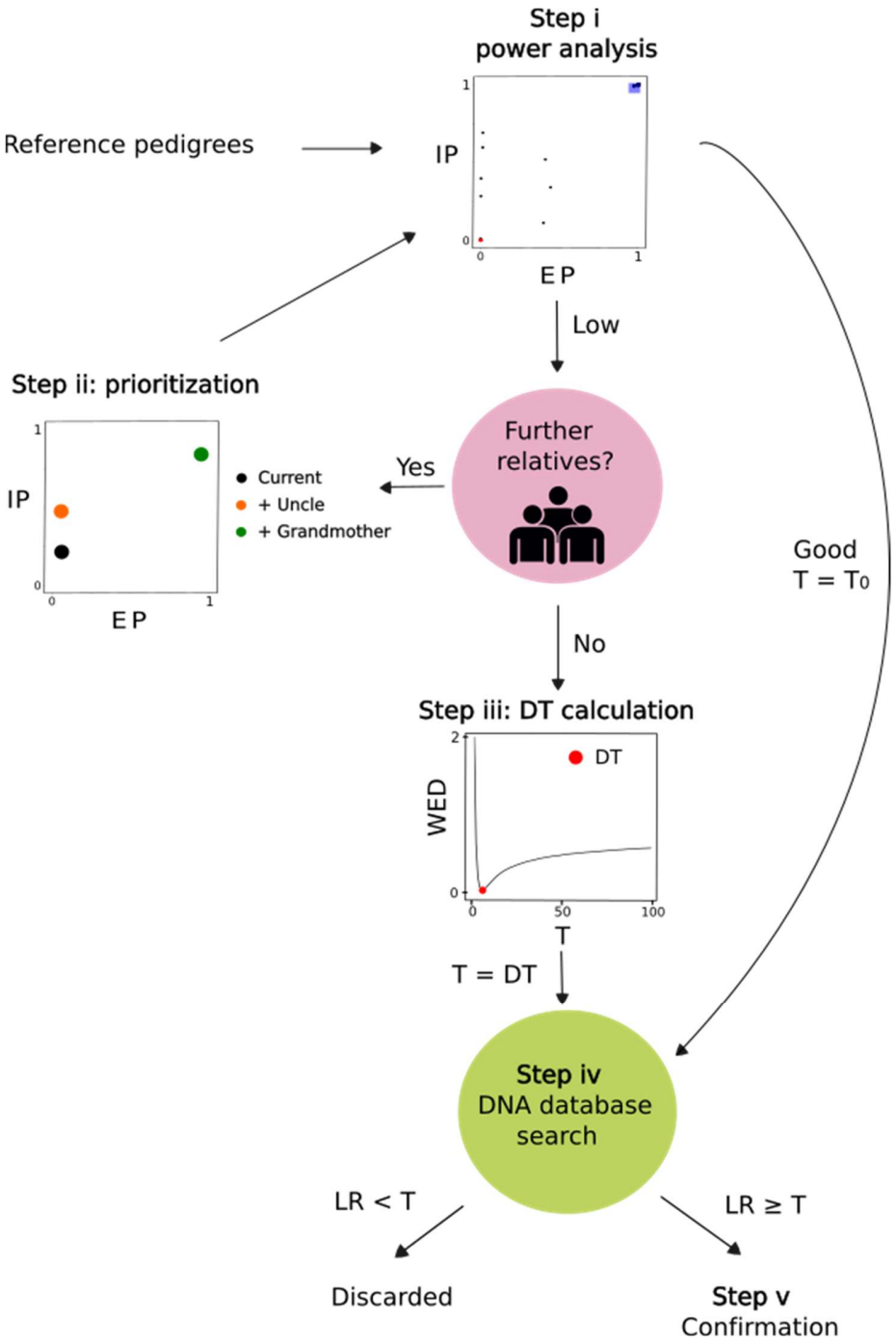

Figure 3: General strategy for DVI/MPI large scale cases. See main text for a detailed description.

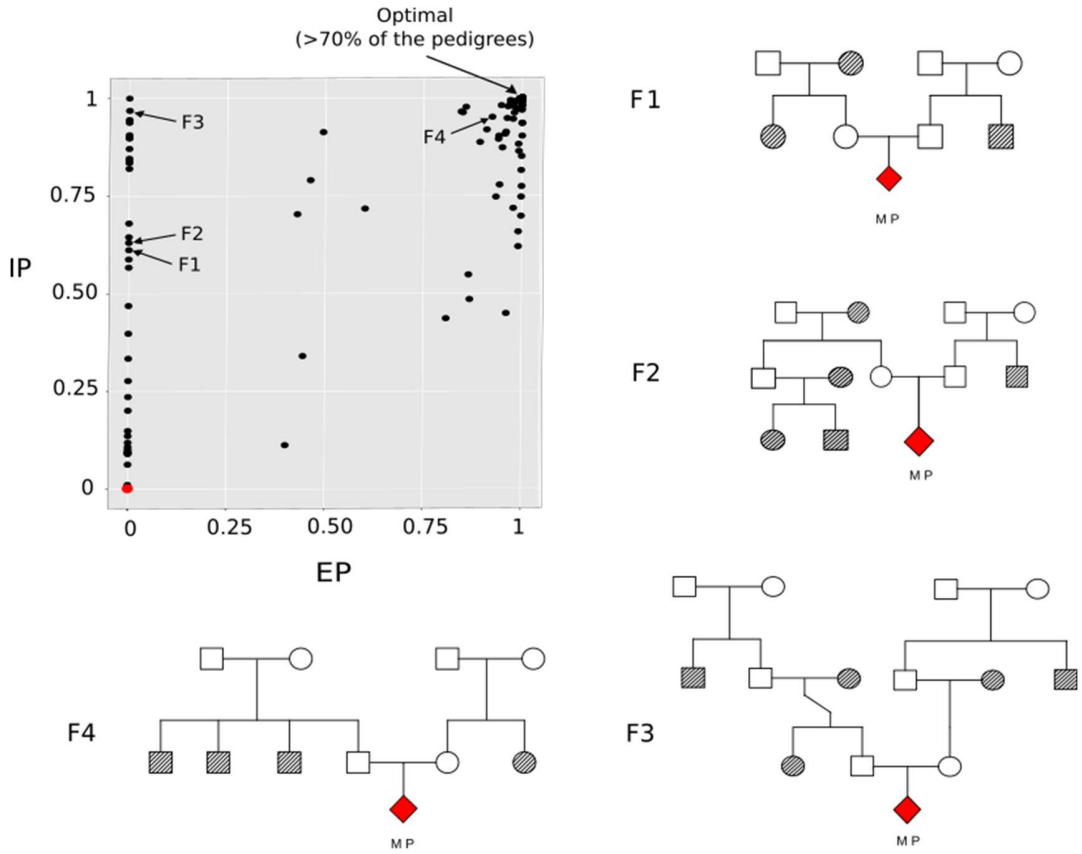

Figure 4: DNA-based identification capability of the pedigrees present in the BNDG. Power plot describ- ing the current statistical power (with 23 markers). Each point represents the EP and IP for a specific pedigree. Example pedigrees are shown.

database) significantly increases as a result of increasing from 15 to 23 markers. In contrast, for median EP there was a smaller increment. 2nd and 3rd degree relatives to the missing person are very common in pedigrees present in the BNDG. Some of those pedigrees may give high LR values resulting from sharing low frequency alleles with POI, but there is also a risk of false positives. The impact of incorporating more genetic markers on $IP_{10,000}$ and EP certainly depends on the pedigree structure (supplementary Ta- ble 1). The addition of more markers does not improve EP for pedigrees with too few or distant relatives to enable exclusion. For pedigrees F1, F2 and F3 there are no impact on the EP when the new markers were added. Considering the $IP_{10,000}$, F1 and F2 have comparable values. Despite not having the genetic

profile of the paternal uncle, the incorporation of the profiles of two paternal cousins and their mother was enough to obtain an acceptable $IP_{10,000}$. F3 shows a high IP with 23 markers but the EP is equal to zero.

Likelihood ratio distribution analysis. Firstly we analyzed the LR distributions obtained from conditional simulations for the pedigrees with low statistical power (some examples are shown in Figure 5). Pedigrees F5 and F10 have very low $IP_{10,000}$ and it is evident that the default LR threshold of 10,000 will lead to a high FNR. The remaining pedigrees (F6 to F9) have $IP_{10,000} > 0$, topped by F8 ($IP_{10,000}$ = 0.62). Interestingly, the probability of finding an unrelated person exceeding the LR threshold is negligible in all cases, demonstrating that T = 10,000 is sufficient for avoiding false positives.

LR Decision Threshold. Curves shown in Figure 6 were obtained for pedigrees F5 to F10 , allowing us to define a DT using the previously illustrated criterion (see Figure 2). Figure 6 shows the FPR-FNR curves and optimal points determining DT for the low-powered pedigrees F5-F10. Numerical results are given in Supplementary Table 2. For example, we see that F7 gave DT = 10, with FNR = 0.03 and FPR = 0.002. In a database of ten thousand people, this corresponds to a manageable number of 20 expected false positives. In contrast, F6 has lower DT and more false positives are expected (around 120 in a database of ten thousand). The most complex cases are pedigrees F5 and F10 where DT is very low (equal to 6) and a high number false positives are expected (more than 200). Using F5 as an example, Figure 7A shows WED as a function of T , elucidating how $T < DT$ rapidly increases WED because of the FPR increment. In contrast, when $T > DT$, WED increases with a lesser slope. Figure 7B shows how DT optimization improves IP from $IP_{10,000}$ = 0.01 to IP6 = 0.47, at the cost of increasing FPR from 0 to 0.021.

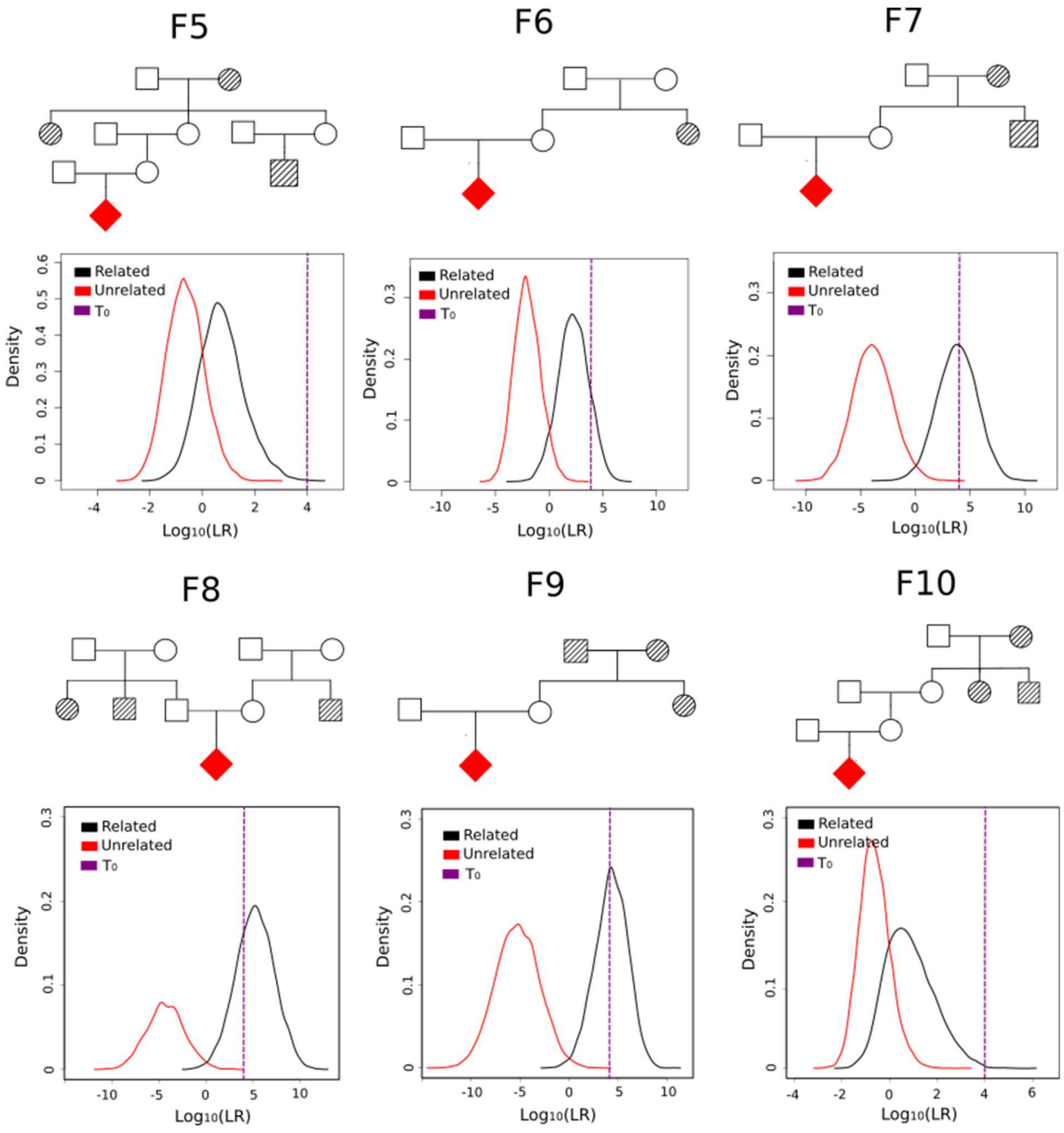

Figure 5: LR distribution of six pedigrees with low statistical power. Log_{10}(LR) distributions for both related (POI is MP) and unrelated (POI is a random person from the population) are plotted. The vertical purple lines indicate a LR threshold T_0 = 10,000. The axes are scaled according to the obtained values.

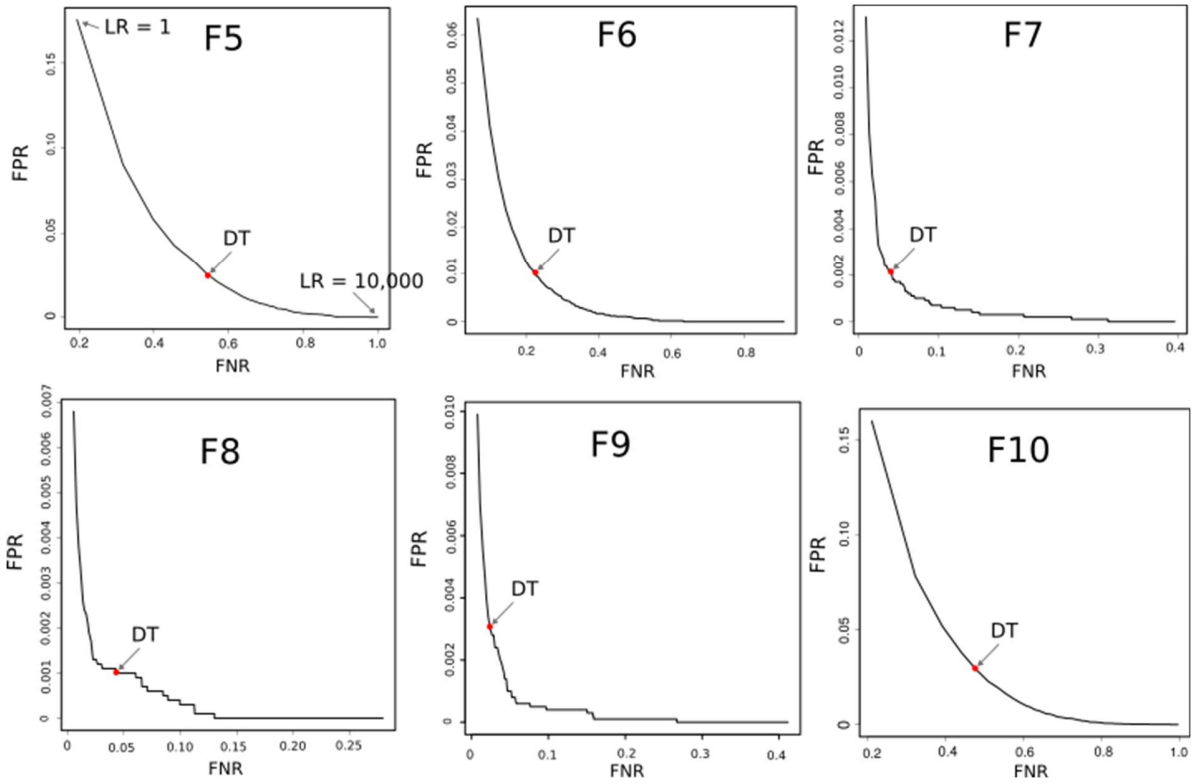

Figure 6: Curves of false positive rates (FPR) and false negative rates (FNR) of the six pedigrees analyzed in Figure 5.

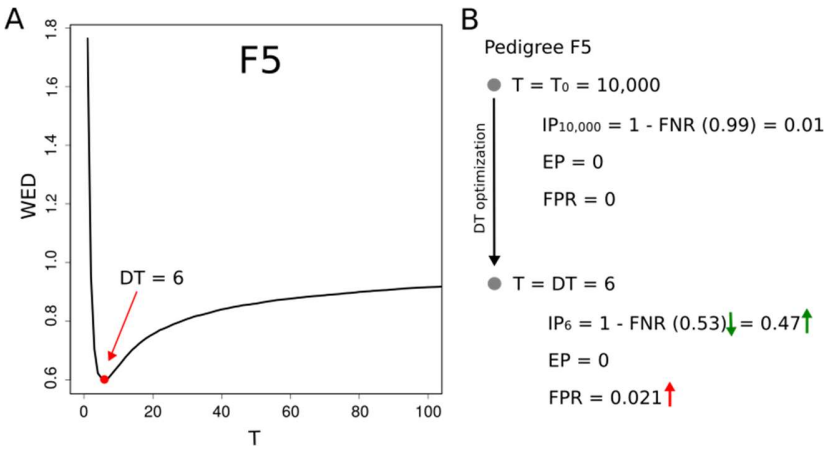

Figure 7: A. Weighted Euclidean Distance (WED) for pedigree F5. B. DT, FPR and FNR for F5 if a T_0 = 10,000 or DT = 6 are considered.

Comparison with other threshold selection strategies. Finally, we compare the DT approach with other methods, some of them reviewed by Kruijver et al [8]. The alternative strategies analyzed here are:

(i) FPR fixed: when FPR is selected for all pedigrees based on laboratory resources, so LR threshold and FNR vary in each one;

(ii) FNR fixed: when FNR is selected, therefore FPR and LR threshold vary;

(iii) LR threshold fixed: when the same LR threshold is selected for all reference pedigrees, named T_l (lower threshold), and it is used as a cut-off for gathering more genetic analysis. In this last case FNR and FPR vary for each reference pedigree. For example, we choose arbitrarily: FPR = 0.01 for (i); FNR = 0.03 for (ii); and T_l = 10 for (iii). Results are shown in Supplementary Table 3. We can observe differences between the alternative strategies and the proposed method. Underpowered pedigrees (F5 and F10) show more divergence between fixed methods and the DT strategy. Particularly, fixed FNR gives useless thresholds for F5 and F10 with FPR = 1. Also, FPR fixed gives higher FNR than DT approach in F5, F6, F8, and F10. In F7 and F9 a slightly larger FNR obtained with DT allows much lower FPR. Fixed T_l gives higher FNR in all cases except F7, where DT = T_l = 10.

Discussion

In this paper we propose a general strategy for dealing with MPI/DVI cases in large-scale DNA database searches. Particularly we provide a statistical method for decision-making for cases with low statistical power. Bayesian decision theory has been proposed as a way of minimizing the risk of taking decisions [19]. It requires some cost to be specified for false positives and false negatives and also the use of priors. This approach is useful in DNA laboratories that treat paternity cases and, as explained by Hedell et al. [20], could be applied also to other problems in forensic genetics. There are several studies applying Bayesian network and decision theory approaches in order to make decisions in database searching [21,22].

In our case, we propose a strategy for taking advantage of the available genetic information from those pedigrees with low statistical power, that despite being insufficient for a conclusion, could be used to select a set of POIs that gives a LR ≥ DT for incorporating more genetic analysis in order to resolve the cases. Furthermore, the approach proposed by Tillmar et al [19] provides a method for decision making when different sets of markers could be incorporated. Comparisons between sets of markers are based on a function relating the expected additional value of information from new data to the amount of information already obtained from initial data. A low DT will give more false positives implying more cases where additional genetic analyses must be performed. On the good side, it will have a lower FNR and decrease the probability of missing an identification in the database search.

In most cases with low statistical power, we were able to obtain an useful DT in pedigrees F6, F7, F8 and F10 (Figure 6). DT approach gives a manageable FPR whilst reducing FNR. For families F5 and F10, DT optimization results in a high FNR and FPR (Figure 6). Despite this, the obtained DT is better than the original T_0 = 10,000, as shown in Figure 7. Incorporating more information from other DNA markers or preliminary investigation could help in these complex cases to reduce the expected number of false positives. Importantly, this would allow to decrease the weight of the FPR in WED calculations (equation 2) leading to a lower DT and therefore a lower FNR.

Comparisons with alternative strategies (Supplementary Table 3) show how DT approach leads to useful thresholds in all cases while other approaches not, for example when FNR is fixed with F5 and F10. When database search is carried out with many reference pedigrees, looking for a useful FPR, FNR or LR cut-off for all cases could be difficult. DT approach allows dealing with underpowered pedigrees obtaining thresholds with manageable FNR and FPR. In well-powered cases an $ad-hoc$ criteria with T_0 = 10,000 is enough to guarantee avoiding false positives without risk of false negatives, as proposed in different contexts by Ge at al [23], Kling et al [1] and revised by Kruijver et al [8].

Generally speaking, a threshold will be referred to as optimal if it classifies most of the individuals correctly when accounting for the relative costs of errors as explained in the definition of WED. Time and economic resources for gathering more genetic information may be limited in certain contexts. Large scale cases hampered by low statistical power lead to many false positives. So, forensic scientists have to decide about how many resources are willing to put in order to avoid missing an identification. All this information results in giving some differential cost between the FPR and FNR. For example, with unlimited economic resources for gathering more genetic information the cost of FPR will be low (so W1 < W2) and it will lead to lower DT. In contrast, if resources are limited the cost will be higher (W1 > W2) and higher DT will be obtained. The decision of cost selection should be done by the forensic scientist and, despite being $ad-hoc$, contributes to mathematically formalize the identification process. In this study we have only considered autosomal STR markers, but many forensic labs incorporate other markers in addition. For example, mitochondrial DNA and Y-STRs markers could be used to reduce the expected number of false positives. For instance, for pedigree F5, all potential identifications (with LR ≥ DT) that do not share mitochondrial DNA with the great-grandmother could be excluded. The same thing, but only applicable to males, applies to Y-STRs. As an example we can analyze pedigree F1 where a paternal uncle is available so all male POIs that do not share Y-STRs with the reference individual, could be excluded. All these strategies are powerful and have the advantage of only requiring genotyping one of the members of the pedigree. They could be used as a filter in database search, its success will depend on the frequency of the haplotypes in mtDNA or alleles in Y-STRs. Nevertheless, it does not avoid genotyping POIs that, with a high FPR, could be expensive in terms of time and resources. Finally, in the case of BNDG, baby theft or child trafficking, information such as age, geography and birth date, could complement the DNA-based identification. Methods for formalizing and incorporating this type of data in database searches have been suggested [24,25].

Conclusion

In this paper we have described a method for dealing with low statistical power in DNA-based identification through database search. Moreover, we have exemplified the strategy using real cases of Missing Grandchildren of Argentina. The algorithm and functions for the calculations of FNR, FPR and DT are implemented in FOGSA, which is freely available from https://github.com/MarsicoFL/FOGSA.

Acknowledges

We thank BNDG members for fruitful discussions.

Author contributions

FLM: Conceptualisation, Methodology, Visualisation, Data curation, Resources, Software, Writing - original draft, Project administration. **MDV**: Conceptualisation, Methodology, Visualisation, Data curation, Software, Writing - review, Writing - original draft. & editing. **TE**: Conceptualisation, Methodology, Visualisation, Data curation, Resources, Software, Writing - review, Writing - original draft. & editing. **MHP**: Conceptualisation, Methodology, Visualisation, Data curation, Resources, Writing - review, Writing - original draft. & editing.

Declaration of competing interests

None.

References

1. Kling, D., Egeland, T., Pin˜ero, M. H., Vigeland, M. D. (2017). Evaluating the statistical power of DNA-based identification, exemplified by 'The Missing Grandchildren of Argentina'. Forensic Science International: Genetics, 31, 57-66.

2. Ge, J., Budowle, B., Chakraborty, R. (2011). Choosing relatives for DNA identification of missing persons. Journal of Forensic Sciences, 56, S23-S28.

3. Baeta, M., Nu´n˜ez, C., Cardoso, S., Palencia-Madrid, L., Herrasti, L., Etxeberria, F., de Pancorbo, M. M. (2015). Digging up the recent Spanish memory: genetic identification of human remains from mass graves of the Spanish Civil War and posterior dictatorship. Forensic Science International: Genetics, 19, 272-279.

4. Vigeland, M. D., Marsico, F. L., Herrera Pin˜ero, M., Egeland, T. (2020). Prioritising family members for genotyping in missing person cases: A general approach combining the statistical power of exclusion and inclusion. Forensic Science International: Genetics, 49, 102376. https://doi.org/10.1016/j.fsigen.2020.102376.

5. Slooten, K. (2020). Likelihood ratio distributions and the (ir) relevance of error rates. Forensic Science International: Genetics, 44, 102173.

6. Cho, S., Shin, E. S., Yu, H. J., Lee, J. H., Seo, H. J., Kim, M. Y., Lee, S. D. (2017). Set up of cutoff thresholds for kinship determination using SNP loci. Forensic Science International: Genetics, 29, 1-8.

7. Li, L., Ye, Y., Song, F., Wang, Z., Hou, Y. (2019). Genetic structure and forensic parameters of 30 InDels for human identification purposes in 10 Tibetan populations of China. Forensic Science International: Genetics, 40, e219-e227.9

8. Kruijver, M., Meester, R., Slooten, K. (2014). Optimal strategies for familial searching. Forensic Science International: Genetics, 13, 90-103.

9. Penchaszadeh, V. B. (1997). Genetic identification of children of the disappeared in Argentina. Journal of the American Medical Women's Association , 52(1), 16-21.

10. Berra, J., Grinspon, D., Liwski, N., Binz, M. T. (1986). Genetical identification of "missing" children in Argentina. In 11th Congress of the Society for Forensic Haemogenetics (Gesellschaft Fu¨r Forensische Blutgruppenkunde eV) (pp. 443-448). Springer, Berlin, Heidelberg.

11. Rittner, C., Schneider, P. M. (Eds.). (2013). Advances in Forensic Haemogenetics: 14th Congress of the International Society for Forensic Haemogenetics (Internationale Gesellschaft for forensische Ha¨mogenetik eV), Mainz, September 18–21, 1991 (Vol. 4). Springer Science Business Media.

12. Penchaszadeh, V. B. (2001). Use of DNA Identification in Human Rights Work to Reunite Families in Latin America. eLife Sciences, 1-8.

13. Egeland, T., Pinto, N., Vigeland, M. D. (2014). A general approach to power calculation for relationship testing. Forensic Science International: Genetics, 9, 186-190.

14. Ricciardi, F., Slooten, K. (2014). Mutation models for DVI analysis. Forensic Science International: Genetics, 11, 85-95.

15. Slooten, K. J., Egeland, T. (2016). Exclusion probabilities and likelihood ratios with applications to mixtures. International Journal of Legal Medicine, 130(1), 39-57.

16. Rota, M., and Antolini, L. (2014). Finding the optimal cut-point for Gaussian and Gamma distributed biomarkers. Computational Statistics and Data Analysis, 69, 1-14.

17. Egeland, T., Mostad, P. F., Meva˚g, B., Stenersen, M. (2000). Beyond traditional paternity and identification cases: selecting the most probable pedigree. Forensic Science International, 110(1), 47-59.

18. Kling, D., Tillmar, A. O., Egeland, T. (2014). Familias 3–Extensions and new functionality. Foren- sic Science International: Genetics, 13, 121-127.

19. Tillmar, A. O., Mostad, P. (2014). Choosing supplementary markers in forensic casework. Forensic Science International: Genetics, 13, 128-133.

20. Hedell, R., Hedman, J., Mostad, P. (2018). Determining the optimal forensic DNA analysis procedure following investigation of sample quality. International Journal of Legal Medicine, 132(4), 955-966.

21. Gittelson, S., Biedermann, A., Bozza, S., and Taroni, F. (2012). The database search problem: A question of rational decision making. Forensic science international, 222(1-3), 186-199.

22. Biedermann, A., Taroni, F. (2012). Bayesian networks for evaluating forensic DNA profiling

evidence: a review and guide to literature. Forensic Science International: Genetics, 6(2), 147-157.

23. Ge, J., Chakraborty, R., Eisenberg, A., Budowle, B. (2011). Comparisons of familial DNA database searching strategies. Journal of forensic sciences, 56(6), 1448-1456.

24. Caridi, I., Dorso, C. O., Gallo, P., Somigliana, C. (2011). A framework to approach problems of forensic anthropology using complex networks. Physica A: Statistical Mechanics and its Applications, 390(9), 1662-1676.

25. Caridi, I., Alvarez, E. E., Somigliana, C., Puerto, M. S. (2020). Using already-solved cases of a mass disaster event for prioritizing the search among remaining victims: a Bayesian approach. Scientific Reports, 10(1), 1-11.

Supplementary

	IP		EP	
Nº markers	15	23	15	23
Median	0.81	0.92	0.84	0.87
F1	0.31	0.57	0	0
F2	0.22	0.63	0	0
F3	0.41	0.95	0	0
F4	0.37	0.91	0	0.89
F5	0	0.01	0	0
F6	0.02	0.23	0	0
F7	0.24	0.54	0	0
F8	0.31	0.72	0	0
F9	0.27	0.56	0.52	0.86
F10	0	0	0	0

Supplementary table 1 - $IP_{10,000}$ and EP with 15 and 23 Autosomal STRs markers for the median of all reference pedigrees analyzed in the BNDG reference database.

Pedigree	DT	FPR	FNR
F5	6	0.021	0.53
F6	8	0.012	0.21
F7	10	0.002	0.03
F8	12	0.001	0.02
F9	7	0.003	0.02
F10	6	0.024	0.47

Supplementary table 2 - LR decision threshold (DT), false positive rate (FPR) and false negative rate (FNR) when T= DT, for pedigrees F5 to F10.

	$T_l = 10$		FPR = 0.01		FNR = 0.03	
Pedigree	FPR	FNR	T_l	FNR	T_l	FPR
F5	0.011	0.61	11	0.64	0 (<0.1)	1
F6	0.010	0.23	11	0.22	24	0.005
F7	0.002	0.03	1	0.01	10	0.002
F8	0	0.03	1	0.06	12	0
F9	0.001	0.04	3	0.01	9	0.002
F10	0.014	0.64	12	0.67	0 (<0.1)	1

Supplementary table 3 - Error rates and thresholds for alternative strategies. When FNR and FPR were fixed, closest LR integer was choosen.