Forensic Science International: Genetics 49 (2020) 102376

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Prioritising family members for genotyping in missing person cases: A general approach combining the statistical power of exclusion and inclusion

Magnus D. Vigeland [a,*], Franco L. Marsico [b], Mariana Herrera Piñero [b], Thore Egeland [c]

[a] Department of Medical Genetics, University of Oslo, PB 4956 Nydalen, 0424 Oslo, Norway
[b] Banco Nacional de Datos Genéticos, Avda. Córdoba 831, C1054AAH CABA, Argentina
[c] Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Chr. M. Falsens vei 1, 1433 Aas, Norway

ARTICLE INFO

Keywords:
Missing person identification
Power analysis
Inclusion power
Exclusion power
Kinship testing
Likelihood ratio
Missing grandchildren of Argentina
Conditional marker simulation

ABSTRACT

Missing person identification typically involves genetic matching of a person of interest against relatives of the missing person. In cases with few available relatives, exhumations or other substantial efforts may be necessary in order to secure adequate statistical power. We propose a simulation approach for solving prioritisation problems arising in such cases. Conditioning on the already typed individuals we estimate the power of each alternative, both to detect the true person, and to exclude false candidates. Graphical summaries of the simulations are given in complementary *power plots*, facilitating interpretation and decision making. Through a series of examples originating from the well-known *Missing grandchildren of Argentina* we demonstrate that our method may untangle complex prioritisation problems and other power-related questions. In particular we offer novel insights in recent cases where only children of the potential match are available for testing. We also show that X-chromosomal markers may give high statistical power in missing person identification, but that this requires careful selection of relatives for genotyping. All simulations, power calculations and plots are done with the R package `forrel`.

1. Introduction

Missing person identification cases are an important subclass of kinship testing problems in forensic genetics. A typical scenario involves a single missing person (MP) for which no genetic data is available. Relatives of MP are typed with a battery of genetic markers, ready to be matched against any person of interest (POI). The evidence is usually measured by the likelihood ratio (LR) comparing the hypothesis that POI is MP to the hypothesis that POI is unrelated to the family [1].

In the context of missing person identification, the *statistical power* concerns the information content of the collected reference data. Loosely speaking, the power can be thought of as probability of reaching a reliable conclusion whenever a POI is genotyped and matched. If this is deemed too low, further data must be gathered, for instance by recruiting additional family members. A central point is that conclusions in both directions are of interest: The *inclusion power* (IP) is the probability that the likelihood ratio will exceed a prescribed threshold if POI really is MP [2,3], while the *exclusion power* (EP) is the probability that a random unrelated individual can be excluded on the grounds of genetical inconsistencies with the family members [2,4]. A unified approach for power analysis in missing person cases, combining IP and EP, was introduced by Kling et al. [2], and has been adopted by subsequent authors [5].

Our primary aim is to provide a practical way to solve *prioritisation problems* in underpowered missing person cases. Above all we seek robust answers to the question: *Which additional relatives should be genotyped?* Previous authors have scrutinised special cases of this problem, using genotype simulation to estimate the distribution of LR [6,3]. Our work differs from these studies in several ways, both in purpose and scope. Firstly, where previous works tend to focus only on inclusion, our context calls for a combination of IP and EP. Secondly, given the heterogeneous nature of missing person cases, our main goal is to provide a general method, rather than explicit results in selected cases. Finally, we advocate the use of *conditional simulations* rather than unconditional

* Corresponding author.
E-mail addresses: magnusdv@medisin.uio.no (M.D. Vigeland), franco.lmarsico@gmail.com (F.L. Marsico), mherrerapinero@gmail.com (M. Herrera Piñero), thore.egeland@nmbu.no (T. Egeland).

https://doi.org/10.1016/j.fsigen.2020.102376
Received 13 May 2020; Received in revised form 11 August 2020; Accepted 17 August 2020
Available online 11 September 2020

gene dropping [2]. For instance, if the parents of MP have genotypes A/A and A/B for a certain marker, the conditional distribution for MP has 50% probability for each of A/A and A/B. In contrast, unconditional simulation proceeds by disregarding the known data and drawing random founder genotypes in each simulation.

Unconditional simulations are useful to investigate generic cases without available data. For instance, the prospects of distinguishing full sibs from half sibs, with a specified marker set, can be resolved using unconditional simulations. In the present work, however, we address cases where some reference individuals are already genotyped. By conditioning on the already known data our conclusions become case-specific rather than averaged over pedigrees with similar structure.

It is an inescapable fact that no single measure adequately captures all power aspects of missing person cases. As a practical response to this we introduce *power plots* providing graphical summaries of the simulation results. We find these plots to be effective tools for comparing the strength of alternative options, for example when prioritising family members for additional genotyping. Importantly, the power plots are easily interpretable also for decision makers with limited understanding of the technical concepts involved in statistical power analysis.

The famous set of missing person cases commonly referred to as the *Missing grandchildren of Argentina* serves as the main motivation behind this work. Several hundred children abducted during the military dictatorship 1976–1983 are still unaccounted for, and continuously searched for by the Banco National de Datos Geneticos (BNDG) in Buenos Aires. Kling et al. [2] concluded that approximately a third of the unsolved cases in BNDG were underpowered at the time of writing. Massive efforts are undertaken in order to improve the information content in these families, including large-scale exhumations of diseased relatives, and locating long-time emigrants. The examples we present are based on current unsolved cases provided by BNDG. For historical background about the missing grandchildren of Argentina and similar cases from other Latin American countries, we refer to [7–9]. Large-scale cases of the same kind from other parts of the world include those reported in [5,10]. Also related to our work is the recent paper by Caridi et al. [11], using Bayesian methods to prioritise identifications among victims of the Argentinian military dictatorship. A recent phenomenon in the search for the Argentinian grandchildren is the emergence of *missing great grandchildren*. With an increasing rate, young people request testing at BNDG, believing that one of their parents is among the missing grandchildren. Apart from the ethical and legal aspects of this development, it also warrants a re-analysis of the entire reference database in terms of power. Intuitively, it is more difficult to match a great grandchild than a grandchild. But how large is this power reduction in terms of IP and EP? And how much power is *gained* if the alleged great grandchild is accompanied by siblings or other family members? Our methods can be used to answer both of these questions. An example is given in Section 3.3, where we analyse a case where four siblings wonder if their late mother was one of the missing grandchildren.

Other than recruiting additional family members, a common way to increase the power of kinship testing is to expand the set of genetic markers. Traditionally, forensic marker batteries are evaluated and compared using template cases like paternity cases with and without the mother genotyped [12]. At BNDG many families were originally typed with 15 autosomal markers, which is now recognised to be insufficient in many cases. The current standard uses 23 markers, but this may be increased to 33 markers in special cases. In Section 3.4 we illustrate the effect on IP and EP of such marker set expansions.

Another interesting recent development is the increasing use of X-chromosomal markers in kinship testing [13]. This motivates our example in Section 3.5, in which we perform a prioritisation analysis using 12 markers on X.

Whole-genome sequencing is becoming increasingly relevant in forensic genetics [14]. This development is likely to influence also the applications discussed in this paper, but in a statistical framework based on genomic identity-by-descent rather than likelihood ratios as we use here. However, simulation studies for power analysis remain relevant also in forensic genealogy, and tools like the R package `ibdsim2` are available for this purpose (https://CRAN.R-project.org/package=ibdsim2).

The field of forensic kinship testing includes several applications closely related to those addressed in this paper. The flexibility of our simulation approach makes it easily adaptable to power analysis in more general cases of family reunification [15], immigration cases [16] and disaster victim identification [17].

All simulations and power analyses presented in this paper were performed in R package `forrel` [18]. This package offers a variety of forensic pedigree analyses, and is available from the official CRAN repository (https://CRAN.R-project.org/package=forrel). The package documentation includes tutorials for power analysis and further examples.

2. Methods

2.1. Definitions and notation

Fig. 1 illustrates the general setup of a missing person case as studied in this paper. We proceed to give precise definitions of the key concepts.

Throughout, a *missing person case* refers to a scenario where a given pedigree, called the *reference pedigree*, has a single missing person (MP). A subset of pedigree members are genotyped with a battery of forensic markers, but no genetic data is available from MP. The DNA profiles of the typed relatives constitute the *reference data*, denoted $\mathscr{R}$.

In general we assume all markers to be unlinked and in linkage equilibrium, and with known allele frequencies. A discussion of these assumptions is given in Section 4.1; in particular we argue that useful results can be obtained also without independence. Each marker may be assigned a mutation model, specified as a matrix $M = (m_{ij})$ where m_{ij} is the per-segregation probability of a mutation from allele i to allele j. We assume throughout that the reference data $\mathscr{R}$ is *consistent*, meaning that it has non-zero probability in the reference pedigree. Note that this definition allows mutations within $\mathscr{R}$, if appropriately modelled. Complicating factors like dropout, dropin and genotyping errors are ignored.

When a person of interest (POI) is to be matched against the reference, the procedure is to type POI with the same set of markers, and compute a likelihood ratio (LR) comparing the following hypotheses (see Fig. 1):

- H_1: POI is the missing person,
- H_2: POI is unrelated to the family.The formula for LR is

$$\mathrm{LR} = \frac{P(\mathrm{data}|H_1, \Theta)}{P(\mathrm{data}|H_2, \Theta)}, \tag{1}$$

where *data* refers to $\mathscr{R}$ together with the profile obtained from POI, and Θ contains marker properties and other fixed parameters. Note that consistency of $\mathscr{R}$ implies that the denominator in (1) is non-zero; thus LR is always well-defined. A successful reunion, sometimes called an *inclusion*, is declared if the LR exceeds some fixed (but ultimately ad hoc) threshold t. In our real-life examples we will use $t = 10\,000$. Oppositely, we declare an *exclusion* of POI if $\mathrm{LR} = 0$, i.e., if the genotypes of POI are inconsistent with the claimed relationship.

2.2. Inclusion power

Consider a missing person case with reference data $\mathscr{R}$. We define the *inclusion power* $\mathrm{IP}(\mathcal{R})$, or simply IP if $\mathscr{R}$ is understood from the context, to be the probability of declaring a positive match if POI is in fact the missing person:

$$\mathrm{IP} = \mathrm{IP}(\mathcal{R}) = P(\mathrm{LR} \geq t|H_1, \mathcal{R}). \tag{2}$$

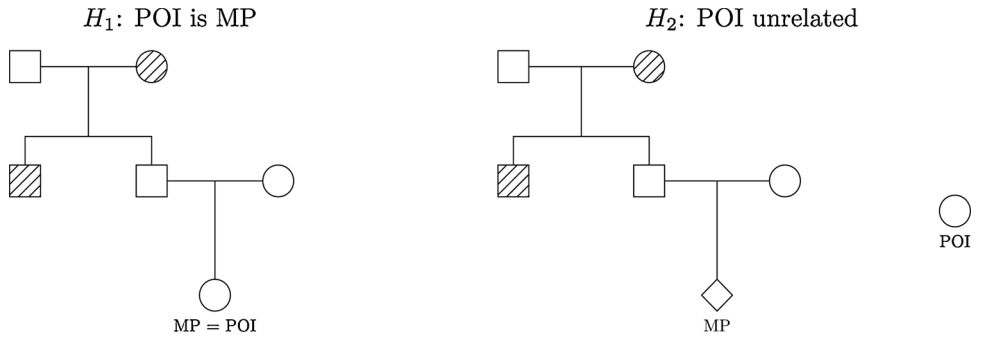

Fig. 1. The competing hypotheses in a typical missing person case. Hatched symbols show typed reference individuals. The sex of the missing person is often unknown. Abbreviations: MP = missing person; POI = person of interest.

Here t denotes the LR threshold, as discussed above. When needed, we include t in the notation as IP_t.

The probability IP_t is intuitive and easy to interpret, especially when t is fixed. However, in more general settings it may be more appropriate to consider other aspects of the LR distribution. In particular we will use the expected value of LR,

$$\mu_{\mathrm{LR}} = \mu_{\mathrm{LR}}(\mathcal{R}) = \mathbb{E}(\mathrm{LR}|H_1, \mathcal{R}), \tag{3}$$

or its logarithm,

$$\lambda_{\mathrm{LR}} = \lambda_{\mathrm{LR}}(\mathcal{R}) = \mathbb{E}(\mu_{\mathrm{LR}}) = \mathbb{E}(log_{10}\mathrm{LR}|H_1, \mathcal{R}). \tag{4}$$

Note that our assumption that $\mathscr{R}$ is consistent guarantees that λ_{LR} is well-defined. We refer to [19,20] for a general discussion of the evaluation of evidence and specific comments on the merits of using a log scale for LR.

In most real-life cases it is infeasible to obtain exact expressions for the distribution of LR, but efficient simulation strategies exist [2,21]. The basic idea is to simulate a large number of DNA profiles for POI under H_1, conditional on $\mathscr{R}$, and to calculate LR in each case by formula (1). As discussed in [2] conditional simulations are preferred over *unconditional* simulations, in which the reference data is simply discarded. Conditional simulations are computationally more expensive, but have the major advantage of enabling unbiased estimates for the specific case in hand, whereas unconditional simulations only allow "averaged" results for similar pedigrees.

2.3. Exclusion power

Again, we consider a missing person case with reference data $\mathscr{R}$. The *exclusion power* $\mathrm{EP} = \mathrm{EP}(\mathcal{R})$ is traditionally defined as the probability that if a random unrelated POI is typed, the result is inconsistent with the reference:

$$\mathrm{EP} = P(\mathrm{atleast1inconsistentmarker}|H_2, \mathcal{R}) \tag{5}$$

Egeland et al. [4] gave a practical formula for EP in general kinship testing, which immediately applies to missing person cases. Briefly, for a single marker s, the exclusion power is seen to be a certain sum of pedigree likelihoods,

$$\mathrm{EP}_s = \sum_g I(g|H_1, \mathcal{R})P(g|H_2, \mathcal{R}), \tag{6}$$

where g denotes a genotype of POI, and $I(g\,|\,H_1)$ is the *inconsistency indicator* with value 1 if g is inconsistent with $\mathscr{R}$ under H_1, and 0 otherwise. Extending to m independent markers gives the total exclusion power

$$\mathrm{EP} = 1 - \prod_1^m (1 - \mathrm{EP}_s). \tag{7}$$

While EP as defined above is simple and mathematically tractable, it is sometimes too limited for practical use. For example, with today's numerous and highly polymorphic forensic markers, a single inconsistent marker may not be considered enough to exclude POI (see also [22]). To overcome this we define a generalised version of EP as follows:

$$\mathrm{EP}^{(k)} = P\big(\mathrm{atleast}\,k\,\mathrm{inconsistentmarkers}|H_2\big) \tag{8}$$

The computation of $\mathrm{EP}^{(k)}$ for $k > 1$ is more complicated than (7), but can still be carried out exactly. See Appendix A.1 for details.

Another problem with the original definition of EP, and indeed $\mathrm{EP}^{(k)}$ for any fixed k, is that it is difficult to appreciate the difference between cases with exclusion probabilities close to either 0 or 1. For example, if two marker sets both have $\mathrm{EP} \approx 1$, it may be hard to justify preferring one over the other. As a better measure of power in such situations we propose to use the *expected number of inconsistent markers*, μ_E. As shown in the Appendix this is readily computed from the marker-wise exclusion probabilities,

$$\mu_E = \sum_{s=1}^m \mathrm{EP}_s. \tag{9}$$

It should be noted that this formula holds irrespective of linkage between the markers.

2.4. Prioritisation problems

A missing person case is *underpowered* if the reference data has low IP or EP. A common strategy to increase the power in such cases is to recruit additional family members for genotyping. Since this may involve substantial efforts, e.g., exhumation of diseased relatives, the question of *who should be prioritised* is essential.

The multifaceted nature of power in missing person cases often makes it difficult to formulate precisely what the prioritisation strategy seeks to optimise. In particular, the relative importance of inclusion and exclusion depends on the application, and may even vary from case to case. Furthermore, in large-scale cases as in Argentina, a balance needs to be striked between the expected number of false positives and false negatives (Marsico et al. *to be submitted*). For these reasons we do not offer a mathematical criterion for optimisation. Instead we propose a two-step simulation procedure which enables detailed comparison of the inclusion and exclusion power in the alternative scenarios.

Procedure 1: Comparing genotyping alternatives in terms of power

Input:

- A missing person case, with baseline reference $\mathcal{R}_0$ containing data from already typed individuals.
- A list of subsets $S_1, \ldots, S_n$ of candidates for additional typing.
- Positive integers p and q, the number of simulations in each step.

Let IP_i denote the expected IP after typing the S_i individuals, for each $i = 1, \ldots, n$, and similarly with EP_i, λ_{LR}^i and μ_E^i.

Output:

- Estimates of IP_i, EP_i, λ_{LR}^i and μ_E^i for each i.
- Sample values of IP, EP, λ_{LR} and μ_E corresponding to p realisations of genotypes for each S_i.
- *Power plots* suitable for comparing the alternatives and drawing conclusions.

Procedure:

(1) For each subset S_i, do:
  (a) Simulate p profiles of the members of S_i conditional on $\mathcal{R}_0$.
  (b) For each $j = 1, \ldots, p$, let $\mathcal{R}_{ij}$ denote the result of the jth simulation, and do:
    (i) Estimate $\text{IP}(\mathcal{R}_{ij})$ and $\lambda_{\text{LR}}(\mathcal{R}_{ij})$ as explained in Section 2.2, using q simulations of MP conditional on $\mathcal{R}_0$ and S_i.
    (ii) Compute $\text{EP}(\mathcal{R}_{ij})$ and $\mu_E(\mathcal{R}_{ij})$ as explained in Section 2.3.
  (c) Estimate IP_i, EP_i, λ_{LR}^i and μ_E^i by averaging over j in the output from part (b).
(2) Plot IP_i against EP_i, $i = 1, \ldots, n$, and similarly λ_{LR}^i against μ_E^i.

A major advantage of the two-step simulation approach described above, is that it not only estimates the expected value IP_i for each i, but in fact the complete distribution of IP in each alternative (and similarly for exclusion). In particular, it should be noted that the estimates of $\text{IP}(\mathcal{R}_{ij}), j = 1, \ldots, p$ obtained in part (1-b-i) are not expected to converge towards IP_i when $q \to \infty$. Rather, their differences reflect the natural genetic variation in the S_i individuals. We often find it useful to visualise this variation by including the estimates of $\text{EP}(\mathcal{R}_{ij})$ and $\text{IP}(\mathcal{R}_{ij})$ in the first power plot, and similarly in the other. Examples of this can be seen in Figs. 2 and 3.

In some cases it may be informative to supplement the above detailed power plots with the empirical marginal LR distributions for each S_i. This is readily done by any good statistics software, using the totality of $p \times q$ values for LR generated in part (1-b-i) of the procedure. See Figs. 4D and 5B for examples.

The notion of exclusion power EP in Procedure 1 may refer to $\text{EP}^{(k)}$ for any fixed k. In the example studies given in the next section we have used the traditional definition with $k = 1$ for simplicity.

3. Results

In this section we present a series of worked examples where we apply Procedure 1, or slight modifications of it, to solve different types of prioritisation problems in missing person cases. The examples are all based on actual unsolved cases from the BNDG database. For anonymity purposes we did not use the true baseline reference data; instead we simulated a set of DNA profiles for the typed individuals, closely matching the format and properties of the true data. In some cases we also modified the pedigree slightly to prevent identification, or to make it more interesting, for instance by introducing additional relatives of MP as candidates for further genotyping.

Unless otherwise specified, all simulations are based on the set of 23 autosomal markers currently employed as standard at BNDG, and a database of Argentinian allele frequencies. This marker set coincides with *Set2* in Section 3.4. The database is available as part of the R package `forrel`. To avoid making the examples overly complex we have not included mutation models. Note however, that the implementation in `forrel` allows mutation modelling both in simulations and power computations.

It is important to bear in mind the effect of conditional simulations in these examples. For example, in Fig. 3 it would be a mistake to assume that the power plots in panels B and C are representative for all reference pedigrees looking like that in panel A. In fact, the plots could change considerably if the genotypes of the reference individuals (hatched) were different. A clear demonstration of this can be seen in our first example, which we give in the next section.

3.1. The importance of conditional simulation

We start with an example demonstrating the importance of conditional simulations in these studies. In the reference pedigree in Fig. 2A, the baseline reference include data from the grandmother (GM) and uncle (U). The power to detect the true MP obviously depends on the actual genotypes of GM and U. If they happen to carry unusually many rare alleles, the inclusion power will be greater than in the opposite case. The white-filled symbols in Fig. 2B and C illustrate this, by showing the power for two different reference data sets for GM and U, denoted $\mathcal{R}_0$ and $\mathcal{R}_0'$. These were picked as the two most diverging among 20 random simulations. As seen in Fig. 2B we have $\text{IP}(\mathcal{R}_0) \approx 0.50$ (white square) while $\text{IP}(\mathcal{R}_0') \approx 0.20$ (white circle).

In order to show how these baseline differences carry over to prioritisation problems, we simulated $p = 10$ profiles for U2 conditional on $\mathcal{R}_0$, and the same for $\mathcal{R}_0'$. We then estimated the power parameters for each of these extended references as in point (1-b) of Procedure 1, with simulation parameter $q = 1000$. As can be seen in Fig. 2B and C, the resulting EP is adequate in either scenario, while IP is consistently above 0.50 in the first case (squares) but below 0.50 in the second case (circles), with averages around 0.70 and 0.30, respectively. We conclude that the addition of U2 would improve the power to a decent level only in one of the two scenarios, namely $\mathcal{R}_0$.

3.2. A larger prioritisation problem

Fig. 3A shows the same reference family as in the previous example, but extended to include further potential candidates for genotyping. As before, the baseline reference data only includes the grandmother and

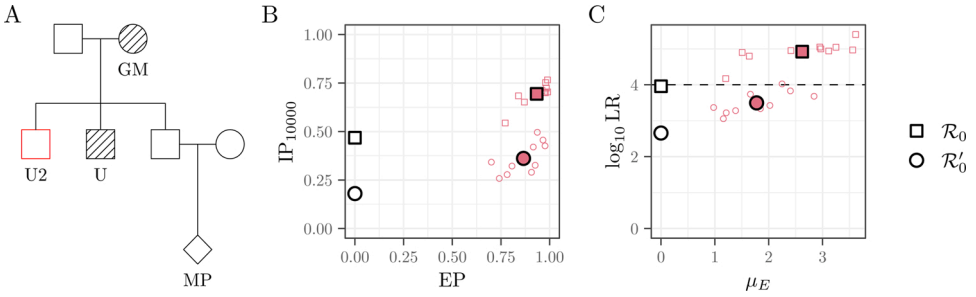

Fig. 2. Power analysis starting from two different reference data sets. (A) The reference pedigree, in which GM and U are typed, and U2 is a potential addition. (B, C) Power plots. The large open symbols show power estimates for two different baseline references, $\mathscr{R}_0$ and $\mathscr{R}_0'$ (see main text). Each minor point corresponds to adding a profile for U2, simulated conditional on $\mathscr{R}_0$ (squares) or $\mathscr{R}_0'$ (circles). Large filled symbols are averages of the minor points of the same shape.

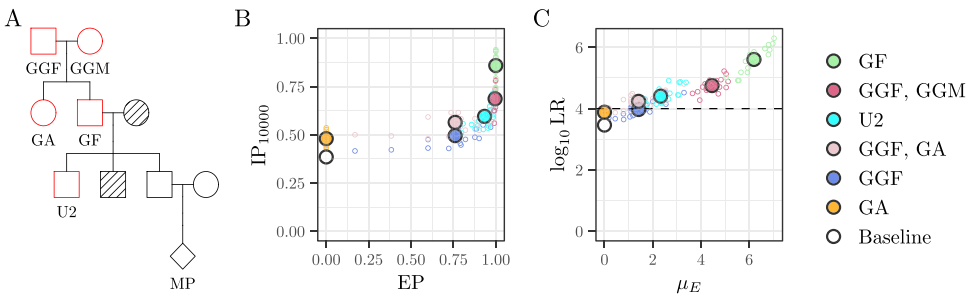

Fig. 3. Simulation results for the prioritisation problem in Section 3.2. (A) The reference pedigree. Hatched individuals are already genotyped; red individuals are candidates for further typing. (B, C) Power plots comparing alternative additions. The open white symbols show baseline power. Coloured symbols represent averages over 20 simulated references. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

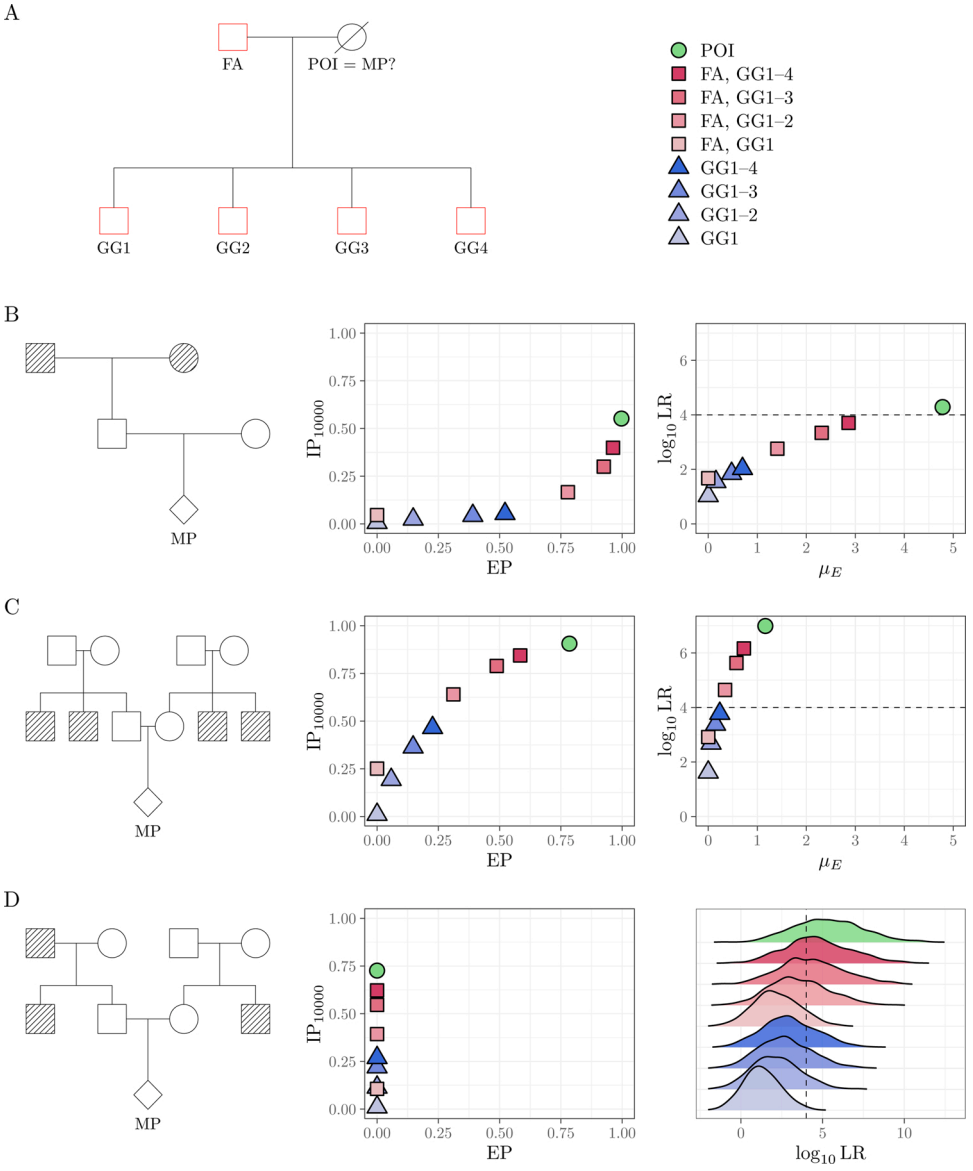

Fig. 4. A power analysis involving missing great grandchildren. (A) The POI family. A female POI is believed to be one of the missing grandchildren. She herself is not available for testing, only her children (GG1–4) and husband (FA). (B–D) Three reference families for which power analysis is performed. In each case, the inclusion and exclusion power is compared when various subsets of the POI family is available for testing. Details about each case are given in the main text.

uncle of MP, but we will now consider five possible additions on the paternal side, MP's grandfather (GF), great grandparents (GGF, GGM), great aunt (GA) and the other uncle (U2).

So who among these should be prioritised? Intuitively it is quite clear that the grandfather (GF) is the most informative, but other aspects of the situation are less obvious:

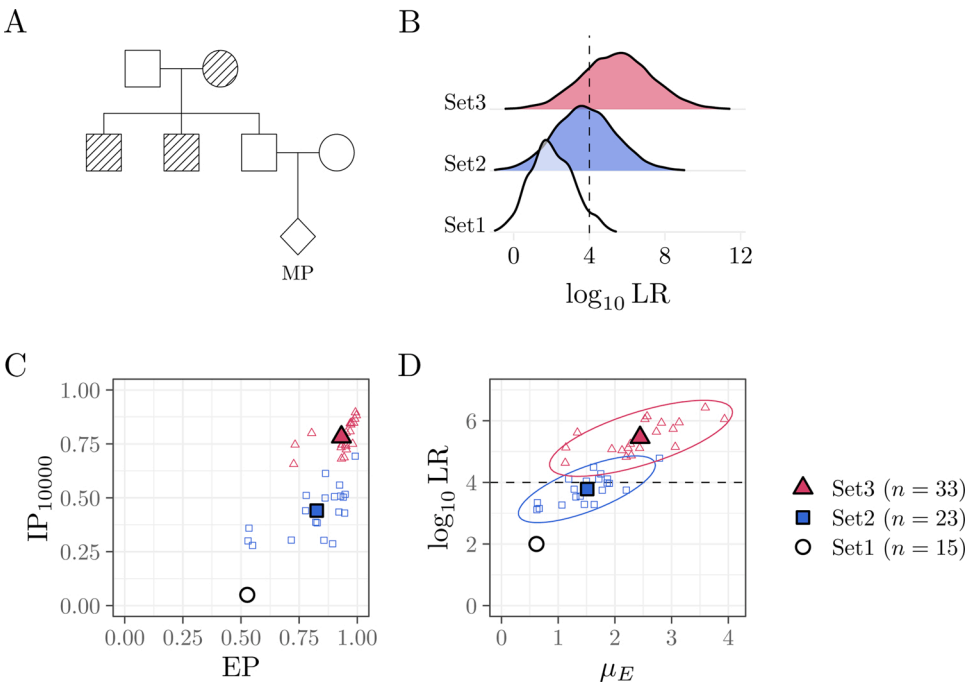

Fig. 5. Analysing the effect of expanding the marker set in a missing person case. The marker sets are described in the main text. (A) The reference pedigree. The hatched individuals are typed with the 15 markers in Set1. (B) LR distributions under H_1. The bottom distribution corresponds to the baseline reference, while the blue and red distributions show the effect of expanding to Set2 or Set3, respectively. (C, D) Power plots. The open circles show baseline power. Each minor square/triangle represents a simulated reference data set augmenting the baseline. Filled symbols show average values. The 95% data ellipses in panel D are produced with default parameters of the function `stat_ellipse()` in the R package ggplot2 [23].

- Are the great grandparents (GGF, GGM) in combination as informative as the grandfather (GF) himself?
- What is the relative importance of the great grandfather (GGF) and the great aunt (GA)?
- How much is gained by including both GGF and GA, instead of just one of them?
- How informative is the uncle U2 compared to the above alternatives?

We also note that with the given baseline reference (hatched individuals), exclusion is impossible since both alleles of MP's father can never be known with certainty. Increasing the exclusion power should therefore be a high priority.

In order to answer the above questions, we followed Procedure 1 with simulation parameters $p = 20$ and $q = 1000$. The results are summarised in the power plots shown in Fig. 3B and C. For example, we see that the grandfather (GF) alone is substantially more informative than including both of his parents (GGF, GGM). In fact, the first power plot indicates that the addition of GF would single-handedly transform this reference family from being seriously underpowered (EP = 0, IP ≈ 0.4), to having good power (EP = 1, IP ≈ 0.87). Furthermore, we see that GGF and GA give very similar increase in inclusion power, but only GGF enables exclusion. Hence if we had to choose between these two, the great grandfather (GGF) would be the correct choice. Interestingly, we note that the combination of GGF and GA does not give much more power than GGF alone. Finally, the extra uncle (U2) comes out only slightly better than the great grandfather (GGF).

3.3. Great grandchildren

The inspiration for this example is a recent case at BNDG, in which four siblings suspect that their late mother was one of the missing grandchildren (Fig. 4A). No data is available from the mother, denoted POI, but all the siblings are willing to donate DNA, and possibly also their father. Fig. 4B–D investigates the power to include/exclude POI through her relatives, in three different reference families. In each case we compared different subsets of POI's family members: Either just a subset of the children (blue triangles), or a set of children accompanied by their father (red squares). We also included the "regular" power (shown by green circles), i.e., when matching directly against POI herself.

The simulation procedure in this example deviated slightly from Procedure 1, since the genotyping alternatives here involve relatives of POI, rather than relatives of MP. In each of the three reference pedigrees (Fig. 4B–D), a baseline dataset $\mathcal{R}_0$ was simulated, containing DNA profiles for the typed individuals (hatched pedigree symbols). Inclusion power estimation was set up by simulating $q = 1000$ sets of profiles for the spouse and children of POI, under the assumption that POI = MP and conditional on $\mathcal{R}_0$. Using this data, 1000 values of $\log_{10}$LR were obtained for each subset of the POI relatives. Similarly, the exclusion power was estimated by simulating 1000 sets of profiles for the POI relatives under H_2, and counting the number of mismatches. Note that exact exclusion powers are computationally out of reach in this example, for reasons outlined in Section 4.3.

Results of the power analyses are shown in the power plots to the right of each pedigree in Fig. 4B–D. We proceed to comment on each case individually, before offering some general remarks.

Fig. 4B. The reference data here only includes the two paternal grandparents of MP. As seen by the green point in the middle panel, this gives good regular exclusion power (EP ≈ 1) and decent inclusion power (IP just over 0.50). However, if we only have data from the children of POI, the power nearly vanishes (blue triangles). The key to securing reasonable power is to recruit their father, i.e., POI's spouse (FA). With data from him and three or four children, much of the regular power is retained. This case also illustrates how the two power plots complement each other: The first plot nicely separates the blue triangles, while the second plot emphasises the difference between the two topmost points.

Fig. 4C. In this reference family two uncles on each side of MP are genotyped. The regular power is good, with EP ≈ 0.80 and IP ≈ 0.90. However, if only GG1 is available, both EP and IP plummet to near 0. As expected, EP and IP increase somewhat when more siblings of GG1 are added, but even with four children the case is severely underpowered (EP ≈ 0.20, IP ≈ 0.50). As in the previous case, the inclusion of the children's father (FA) boosts the power significantly. In particular, we see that FA together with three or more children almost recover the regular power.

Fig. 4D. The final reference family differs from the two others, in that exclusion of POI is not possible. Hence the points in the power plot cluster on the EP = 0 axis. In such cases it is more informative to

compare the distributions of $\log_{10}$LR, as shown in the rightmost plot.

In addition to the case-specific remarks above, some general insights may be drawn from these analyses:

- Matching against children of POI has generally much lower power than matching against POI directly. (Cases where parents of MP are typed are exceptions to this.)
- Recruiting POI's partner boosts the information content of the children, and is crucial for increasing the power.
- If enough close relatives of POI are included, the power may be almost as good as if POI were available.

We emphasise that the analysis in this example, and the above conclusions, apply exclusively to autosomal markers. In actual casework involving great grandchildren, lineage markers (Y-chromosomal or mitochondrial, depending on POI's sex) could be highly informative and should be considered in addition to the autosomal analysis.

3.4. Expanding the marker set

A common strategy for increasing the power in missing person cases is to re-type the reference individuals with more genetic markers. In this process it is valuable to estimate the expected power gain. One might also want to compare different available marker kits.

To illustrate such an analysis, we used the reference pedigree in Fig. 5A in which three individuals (hatched) are typed at baseline with 15 markers. As seen in Fig. 5C this gives poor power (IP ≈ 0.05, EP ≈ 0.50), and our task is to estimate the expected improvement by re-typing the three individuals with more markers.

The original 15 markers, and two possible extension sets, are as follows:

- Set1 (15 markers): CSF1PO, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX, VWA.
- Set2 (23 markers): Set1 + D1S1656, D2S441, D10S1248, D12S391, D22S1045, PENTA_D, PENTA_E, SE33.
- Set3 (33 markers): Set2 + D2S1360, D3S1744, D4S2366, D5S2500, D6S474, D6S1043, D7S1517, D8S1132, D10S2325, D21S2055.

Set1 used to be the standard kit at BNDG, but is now replaced with Set2. The larger Set3 is reserved for special cases.

The simulation strategy for this example broadly followed Procedure 1, except that the simulated individuals stayed the same in all steps, while the marker sets varied. To get started, we simulated a baseline reference $\mathcal{R}_0$ using the markers in Set1. Then, for each of Set2 and Set3 we simulated $p = 20$ references extending $\mathcal{R}_0$. Only the additional markers were simulated; no conditional simulations were needed in this step. For each of the resulting 40 data sets, IP and λ_{LR} were estimated through $q = 1000$ conditional simulations of POI under H_1, and EP and μ_E were computed exactly.

The results are displayed in Fig. 5B–D. As evident e.g. from the data ellipses in panel D, there is a large spread in power among the extended reference data sets. Nevertheless, it is clear that extending to Set3 has a high probability of producing a good-powered reference. Settling for Set2 will *sometimes* yield decent power, but usually not: most realisations (minor blue symbols) have IP < 0.50. As in previous examples, we emphasise that these conclusions are specific to the starting point $\mathcal{R}_0$ and cannot be expected to hold in other cases with the same pedigree.

3.5. Power analysis of X-chromosomal markers

For our final example, we perform a power analysis using the Investigator Argus X-12 kit (Qiagen, Hilden, Germany) containing 12 X-chromosomal STR markers. This kit is commonly used in forensics, and an Argentinian frequency database has recently been published [24]. It is known that both linkage and linkage disequilibrium exist between some of these markers, but we argue that this will not influence the broad conclusions of our analysis. We acknowledge that dependencies between markers may affect some statistics, including IP and EP, but the estimates of μ_{LR} and μ_E remain unbiased.

As our example we use the pedigree shown in Fig. 6. The great potential in using X-chromosomal markers is clearly illustrated by the paternal grandmother (GM). The X-chromosomal relationship between her and her granddaughter (MP) is identical to an autosomal mother-daughter relationship, thus she single-handedly gives very good power. Note also that for this particular relationship, marker linkage does not affect LR calculations, since the father of MP is hemizygous.

For this example, we started with an empty baseline reference, i.e., we did not condition on any already typed family members. In all other respects we followed Procedure 1, with parameters $p = 20$ and $q = 1000$.

Fig. 6 shows that no exclusion power and poor inclusion power is expected if only the grandfather (GF) and one aunt (A1) of MP is typed. All the other alternatives considered enable exclusion, and give some increase in the expected inclusion power. However, only the paternal grandmother (GM) of MP improves power sufficiently: with her we can expect decisive power to recognise the true MP, and on average 6 inconsistent markers in an unrelated POI.

4. Discussion

We have proposed an approach to solving prioritisation problems in missing person cases, by comparing the alternatives in terms of their expected power to detect the true missing person, and to exclude unrelated individuals. Our primary aim was to answer specific questions of the kind "Which additional relatives should be genotyped?", but as illustrated by our examples, the method applies more generally to other similar problems. In particular, we have given a first detailed power analysis of a case involving "missing great grandchildren", a scenario of increasing relevance for the work at BNDG in Argentina. Although more work is needed on this subject, some general patterns already emerge from our results presented in Fig. 4. As expected by intuition the statistical power to identify (or exclude) a great grandchild is substantially lower than for a grandchild. But if several siblings join in, reasonable power may be achieved, especially if accompanied by a parent, i.e., the spouse of the missing grandchild.

We have also exemplified the usefulness of X-chromosomal markers in missing person identification. Particularly in cases with sparse reference data we recommend to always run both autosomal and X-chromosomal prioritisation analysis, also in cases where the missing person has unknown sex. For example, suppose that in a given missing person case a choice must be made of which grandparent to exhume. If there is a decent chance that the missing person is female, it is clear from Fig. 6 that the paternal grandmother *a priori* should be given strong priority. Depending on the available data, she will on average increase both IP and EP significantly using X-chromosomal markers, in addition to her autosomal contribution.

In the next sections we discuss some technical aspects of our methods.

4.1. Assumptions

In most applied areas it is typical that the models used in power calculations are simpler than those used when data has been collected. Sometimes good approximations suffice in the initial phase when power assessments are done. Moreover, if one can show that the simpler model gives *unbiased* estimates this may be sufficient for making decisions, even if the model is inaccurate in other respects. Our analysis of X-chromosomal markers in Section 3.5 gives a good example of this: the estimates of λ_{LR} are unbiased in spite of linkage between the markers.

A strong assumption in most forensic genetic analysis, is that allele frequencies are sufficiently accurate. Toscanini et al. [25] report

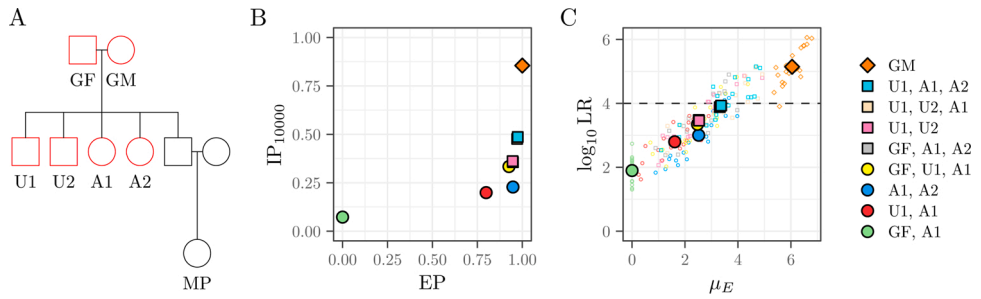

Fig. 6. Power analysis with X-linked markers. (A) Reference pedigree, with no individuals typed from the outset, but six members (red) are candidates for testing with 12 X-chromosomal markers. (B, C) Power plots showing mean estimates for each alternative. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

evidence for population structure and difference in allele frequencies and concludes that "caution should be taken when using a common pooled database with general forensic purposes in Argentina". Obviously, this recommendation is not unique to Argentina. Generally, one is well advised to do calculations with several databases and check if conclusions remain unchanged. We prefer this pragmatic approach to the so-called θ or F_{ST} correction [26,27].

Very rare alleles are particularly challenging, especially when a previously unknown allele is observed in a case. Most forensic calculations require a frequency to be set, giving rise to various ad-hoc rules [28]. Importantly, a single marker may give a very large LR if a rare allele is involved. In the context of power analysis we again recommend pragmatic safe-guarding against unreliable conclusions: If the reference data contains very rare alleles, calculations should be repeated with different reasonable allele frequencies.

4.2. Multiple POIs

Throughout this work we have considered a single POI coming forward. However, the approach may be modified to accommodate several simultaneous POIs, say POI_1, …, POI_r, where we face the hypotheses

$$H_{1,i} : \text{POI}_i = \text{MP}, \quad i = 1, \ldots, r,$$

and H_2: all POIs are unrelated to the family. Recall that in our setting, unlike e.g., in cases of disaster victim identifications, the POIs are not yet genotyped. Hence their LR distributions are identical, and updates of priors or prior odds to the posterior versions only reflect simulation uncertainty. However, the definition of inclusion power (2) may be extended to

$$\text{IP}^* = P(max_i\{\text{LR}_i | H_{1,i}\} \geq t | \mathcal{R}).$$

In this case it would be appropriate to increase the threshold t in order to maintain the same confidence as for a single POI.

4.3. Computational challenges

Markers with many alleles pose a challenge to the exact computation of EP, because of the large number of genotypes we have to sum over in Eq. (6). In standard missing person cases the problem is manageable since only one person is involved: With L alleles, the number of genotypes is then $n(L) = (L+1)L/2$, which is well within the computational limit for any realistic L. For example, the marker SE33 may have as many as $L = 70$ alleles, giving $n(L) = 2485$. However, with two individuals the number of genotype combinations rises to $n(L)^2 = 6,175,225$ which is already on the brink of infeasibility.

There are several ways to overcome this computational challenge. In certain simple cases, exact formulas may be obtained. In more complex cases, we propose simulations, carried out in a manner similar to our estimation of IP. This was the approach taken in our example with great grandchildren in Section 3.3, which required EP calculations involving as many as five individuals.

4.4. Inconsistencies in the reference data

Mendelian inconsistencies are inevitable in case work. Fig. 6 in [2] shows an example where the mother is 7/9.3 and the son 8/8. In this case it is most likely that a mutation occurred, probably the allele 7 was transferred as 8. In cases with inconsistencies for several markers, the explanation may not be mutation but rather a misspecified relationship. For instance, misattributed parentages occur, and are often handled by modifying the pedigree without disclosing the information to the family.

If mutations are deemed to have occurred in the reference data, some practical choices need to be made on how they should be handled. Many of mutation models used in case work dictates all transitions to be possible, with the disadvantage that exclusion becomes trivially impossible. The one-step mutation model, or other models that assign zero probabilities to some transitions [29], *can* give positive exclusion probability. Our practical suggestion is to apply such a model for those markers where the reference data is inconsistent (i.e., a mutation has clearly happened) and disable mutations for the remaining markers.

5. Conclusion

In this paper we have described a strategy for planning additional genotyping in underpowered missing person cases. Furthermore, we have illustrated the method through a series of examples of both practical and theoretical interest, based on true cases of missing grandchildren in Argentina. The simulation algorithm and functions for creating power plots are implemented in the R package `forrel` [18], which is freely available from the official CRAN repository, https://CRAN.R-project.org/package=forrel.

Author contributions

MDV: Conceptualisation, methodology, visualisation, software, writing - original draft, project administration. FLM: Data curation, resources, writing - review & editing. MHP: Data curation, resources, writing - review & editing. TE: Conceptualisation, methodology, software, writing - original draft.

Conflict of interest

The authors declare that there is no conflict of interest.

Appendix A. Computational details

A.1 Exclusion power

In this appendix we produce an exact formula for $\mathrm{EP}^{(k)}$, the probability of at least k inconsistent markers. We also give formulas for the mean and variance of the total number of inconsistent markers.

Let m be the number of markers, and for each $s = 1, \ldots, m$, let I_s be the inconsistency indicator with value 1 if marker s is inconsistent with hypothesis H_1, and 0 otherwise. The exclusion power of marker s is then $\mathrm{EP}_s = \mathbb{E}(I_s)$. Furthermore the total number of inconsistencies, denoted N_E, satisfies

$$N_E = \sum_{s=1}^{m} I_s. \tag{10}$$

If the markers are independent, N_E is a sum of independent Bernoulli variables, implying that it has a Poisson-binomial distribution with point probabilities

$$P(N_E = x) = \sum_{A} \prod_{i \in A} \mathrm{EP}_i \prod_{j \in \overline{A}} (1 - \mathrm{EP}_j), \tag{11}$$

where in the outer sum A runs over all subsets of $\mathscr{M} := \{1, \ldots, m\}$ of size x, and $\overline{A}$ denotes the complement of A in $\mathscr{M}$. As a result, we find that the kth order exclusion power can be computed as

$$\mathrm{EP}^{(k)} = 1 - \sum_{x=0}^{k-1} P(N_E = x) = 1 - \sum_{A} \prod_{i \in A} \mathrm{EP}_i \prod_{j \in \overline{A}} (1 - \mathrm{EP}_j), \tag{12}$$

where A now runs over all subsets of $\mathscr{M}$ of cardinality strictly less than k. (Note that (7) is recovered by setting $k = 1$, so that A is the empty set and $\overline{A} = \mathscr{M}$).

The expected number of inconsistent markers is readily computed from (10). Irrespective of linkage between the markers, we find

$$\mu_E = \mathbb{E}(N_E) = \sum_{s=1}^{m} \mathbb{E}(I_s) = \sum_{s=1}^{m} \mathrm{EP}_s. \tag{13}$$

Finally, if the markers are independent (i.e., unlinked), the variance of N_E is given by

$$\mathrm{Var}\,(N_E) = \sum_{s=1}^{m} \mathrm{EP}_s (1 - \mathrm{EP}_s). \tag{14}$$

References

[1] T.J. Parsons, R.L.M. Huel, DNA and missing persons identification: practice, progress and perspectives.. Handbook of Forensic Genetics. Biodiversity and Heredity in Civil and Criminal Investigation, World Scientific, New Jersey, 2016 https://doi.org/10.1142/9781786340788_0015.

[2] D. Kling, T. Egeland, M.H. Piñero, M.D. Vigeland, Evaluating the statistical power of DNA-based identification, exemplified by 'The missing grandchildren of Argentina', Forensic Sci. Int.: Genet. 31 (2017) 57–66, https://doi.org/10.1016/j.fsigen.2017.08.006.

[3] N. Pinto, R. Simões, A. Amorim, E. Conde-Sousa, Optimizing the information increase through the addition of relatives and genetic markers in identification and kinship cases, Forensic Sci. Int.: Genet. 40 (2019) 210–218, https://doi.org/10.1016/j.fsigen.2019.02.019.

[4] T. Egeland, N. Pinto, M.D. Vigeland, A general approach to power calculation for relationship testing, Forensic Sci. Int.: Genet. 9 (2014) 186–190, https://doi.org/10.1016/j.fsigen.2013.05.001.

[5] K. Yu, W.K. Fung, Evaluation of parentage testing accuracy of child trafficking cases: combining the exclusion probability and likelihood ratio approaches, Forensic Sci. Int.: Genet. 34 (2018) 81–87, https://doi.org/10.1016/j.fsigen.2018.02.002.

[6] J. Ge, B. Budowle, R. Chakraborty, Choosing relatives for DNA identification of missing persons, J. Forensic Sci. 56 (2011) S23–S28, https://doi.org/10.1111/j.1556-4029.2010.01631.x.

[7] A.M. Di Lonardo, P. Darlu, M. Baur, C. Orrego, M.-C. King, Human genetics and human rights. Identifying the families of kidnapped children, Am. J. Forensic Med. Pathol. 5 (4) (1984) 339–347, https://doi.org/10.1097/00000433-198412000-00011.

[8] V.B. Penchaszadeh, Use of DNA identification in human rights work to reunite families in Latin America, eLS (2001) 1–8, https://doi.org/10.1002/9780470015902.a0027009.

[9] S.A. Vishnopolska, A.G. Turjanski, M.H. Piñero, B. Groisman, R. Liascovich, A. Chiesa, M.A. Marti, Genetics and genomic medicine in Argentina, Mol. Genet. Genomic Med. 6 (4) (2018) 481–491, https://doi.org/10.1002/mgg3.455.

[10] M.A. Brandenburg, S.M. Watkins, K.L. Brandenburg, C. Schieche, Operation Child-ID: reunifying children with their legal guardians after Hurricane Katrina, Disasters 31 (3) (2007) 277–287, https://doi.org/10.1111/j.1467-7717.2007.01009.x.

[11] I. Caridi, E.E. Alvarez, C. Somigliana, M.S. Puerto, Using already-solved cases of a mass disaster event for prioritizing the search among remaining victims: a Bayesian approach, Sci. Rep. 10 (1) (2020) 1–11, https://doi.org/10.1038/s41598-020-59841-3.

[12] A.B. Penacino, How many STR markers are enough? Forensic Sci. Int.: Genet. Suppl. Ser. (2019) https://doi.org/10.1016/j.fsigss.2019.10.170.

[13] A.O. Tillmar, D. Kling, J.M. Butler, W. Parson, M. Prinz, P.M. Schneider, T. Egeland, L. Gusmão, DNA Commission of the International Society for Forensic Genetics (ISFG): guidelines on the use of X-STRs in kinship analysis, Forensic Sci. Int.: Genet. 29 (2017) 269–275, https://doi.org/10.1016/j.fsigen.2017.05.005.

[14] A. Tillmar, P. Sjölund, B. Lundqvist, T. Klippmark, C. Älgenäs, H. Green, Whole-genome sequencing of human remains to enable genealogy DNA database searches – a case report, Forensic Sci. Int.: Genet. 46 (2020) 102233, https://doi.org/10.1016/j.fsigen.2020.102233.

[15] M. Baeta, C. Núñez, S. Cardoso, L. Palencia-Madrid, L. Herrasti, F. Etxeberria, M. M. de Pancorbo, Digging up the recent Spanish memory: genetic identification of human remains from mass graves of the Spanish Civil War and posterior dictatorship, Forensic Sci. Int.: Genet. 19 (2015) 272–279, https://doi.org/10.1016/j.fsigen.2015.09.001.

[16] A.O. Karlsson, G. Holmlund, T. Egeland, P. Mostad, DNA-testing for immigration cases: the risk of erroneous conclusions, Forensic Sci. Int. 172 (2–3) (2007) 144–149, https://doi.org/10.1016/j.forsciint.2006.12.015.

[17] C.M. Vullo, M. Romero, L. Catelli, M. Šakić, V.G. Saragoni, M.J.J. Pleguezuelos, C. Romanini, M.J. Anjos Porto, J.P. Prieto, A.B. Castro, A. Hernandez, M.J. Farfán, V. Prieto, D. Alvarez, G. Penacino, S. Zabalza, A.H. Bolaños, I.M. Manterola, L. Prieto, T. Parsons, GHEP-ISFG collaborative simulated exercise for DVI/MPI:

lessons learned about large-scale profile database comparisons, Forensic Sci. Int.: Genet. 21 (2016) 45–53, https://doi.org/10.1016/j.fsigen.2015.11.004.
[18] M.D. Vigeland, T. Egeland, Handling founder inbreeding in forensic kinship analysis, Forensic Sci. Int.: Genet. Suppl. Ser. (2019), https://doi.org/10.1016/j.fsigss.2019.10.175.
[19] I.J. Good, Weight of evidence: a brief survey, Bayesian Stat. 2 (1985) 249–270.
[20] T. Egeland, K. Slooten, The likelihood ratio as a random variable for linked markers in kinship analysis, Int. J. Legal Med. 130 (6) (2016) 1445–1456, https://doi.org/10.1007/s00414-016-1416-2.
[21] M. Kruijver, Efficient computations with the likelihood ratio distribution, Forensic Sci. Int.: Genet. 14 (2015) 116–124, https://doi.org/10.1016/j.fsigen.2014.09.018.
[22] K.J.D. Balloch, J. Marshall, J. Clugston, J.W. Gow, Reporting paternity testing results when 2 exclusions are encountered, Forensic Sci. Int.: Genet. Suppl. Ser. 1 (1) (2008) 492–493, https://doi.org/10.1016/j.fsigss.2007.10.096.
[23] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York, 2016.
[24] M.G. García, C.I. Catanesi, G.A. Penacino, L. Gusmão, N. Pinto, X-chromosome data for 12 STRs: towards an Argentinian database of forensic haplotype frequencies, Forensic Sci. Int.: Genet. 41 (2019) e8–e13, https://doi.org/10.1016/j.fsigen.2019.04.005.
[25] U. Toscanini, L. Gusmão, G. Berardi, A. Amorim, Á. Carracedo, A. Salas, E. Raimondi, Testing for genetic structure in different urban Argentinian populations, Forensic Sci. Int. 165 (1) (2007) 35–40, https://doi.org/10.1016/j.forsciint.2006.02.042.
[26] D.J. Balding, R.A. Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands, Forensic Sci. Int. 64 (2–3) (1994) 125–140, https://doi.org/10.1016/0379-0738(94)90222-4.
[27] J. Buckleton, J. Curran, J. Goudet, D. Taylor, A. Thiery, B.S. Weir, Population-specific FST values for forensic STR markers: a worldwide survey, Forensic Sci. Int.: Genet. 23 (2016) 91–100, https://doi.org/10.1016/j.fsigen.2016.03.004.
[28] B. Martin, B. Ingo, J.M. Butler, R. Fimmers, P. Gill, L. Gusmão, N. Morling, C. Phillips, M. Prinz, P.M. Schneider, W. Parson, Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER), Forensic Sci. Int.: Genet. 24 (2016) 97–102, https://doi.org/10.1016/j.fsigen.2016.06.008.
[29] C.H. Brenner, Multiple mutations, covert mutations and false exclusions in paternity casework, International Congress Series, vol. 1261 (2004) 112–114, https://doi.org/10.1016/S0531-5131(03)01843-0.