Available online at www.sciencedirect.com
ScienceDirect
Forensic Science International: Genetics xxx (2007) xxx–xxx
FSI GENETICS
www.elsevier.com/locate/fsig

Review

ISFG: Recommendations on biostatistics in paternity testing

David W. Gjertson [a], Charles H. Brenner [b], Max P. Baur [c], Angel Carracedo [d], Francois Guidet [e], Juan A. Luque [f], Rüdiger Lessig [g], Wolfgang R. Mayr [h], Vince L. Pascali [i], Mechthild Prinz [j], Peter M. Schneider [k], Niels Morling [l,*]

[a] *Department of Biostatistics, UCLA School of Public Health, Box 951772, Los Angeles, CA 90095-1772, USA*
[b] *6801 Thornhill Drive, Oakland, CA 94611-1336, USA*
[c] *Institute of Medical Biometry, Rheinische Friedrich-Wilhelms-Universität Bonn, Sigmund-Freud-Str. 25, D-53105 Bonn, Germany*
[d] *Institute of Legal Medicine, University of Santiago de Compostela, E-15705 Santiago de Compostela, Spain*
[e] *Institut Genetique Nantes Atlantique, 19 rue Leon Durocher, BP 70425, F-44204 Nantes Cedex 2, France*
[f] *Instituto Nacional de Toxicología y Ciencias Forenses, Servicio de Biología, Ministerio de Justicia. Merced 1, E-08002 Barcelona, Spain*
[g] *Institute of Legal Medicine, University of Leipzig, Johannisallee 28, D-04103 Leipzig, Germany*
[h] *University Clinic for Blood Group Serology and Transfusion Medicine, Medical University of Vienna, A-1090 Vienna, Austria*
[i] *Istituto di medicina legale, Università Cattolica, Largo Francesco Vito, I-100168 Roma, Italy*
[j] *Office of the Chief Medical Examiner, Department of Forensic Biology, 520 First Avenue, New York, NY 10016, USA*
[k] *Institute of Legal Medicine, University of Cologne, D-50823 Köln, Germany*
[l] *Department of Forensic Medicine, Faculty of Health Sciences, University of Copenhagen, 11 Frederik V's Vej, DK-2100 Copenhagen Ø, Denmark*

Received 23 June 2007; accepted 27 June 2007

Abstract

The Paternity Testing Commission (PTC) of the International Society for Forensic Genetics has taken up the task of establishing the biostatistical recommendations in accordance with the ISO 17025 standards and a previous set of ISFG recommendations specific to the genetic investigations in paternity cases. In the initial set, the PTC recommended that biostatistical evaluations of paternity are based on a likelihood ratio principle – yielding the paternity index, PI. Here, we have made five supplementary biostatistical recommendations. The first recommendation clarifies and defines basic concepts of genetic hypotheses and calculation concerns needed to produce valid PIs. The second and third recommendations address issues associated with population genetics (allele probabilities, Y-chromosome markers, mtDNA, and population substructuring) and special circumstances (deficiency/reconstruction and immigration cases), respectively. The fourth recommendation considers strategies regarding genetic evidence against paternity. The fifth recommendation covers necessary documentation, reporting details and assumptions underlying calculations. The PTC strongly suggests that these recommendations should be adopted by all laboratories involved in paternity testing as the basis for their biostatistical analysis.

Keywords: Recommendations; Paternity testing; Immigration testing; Relationship testing; DNA investigations; Genetic testing; Biostatistics; Paternity index (PI); Likelihood ratio; International Society for Forensic Genetics (ISFG); Paternity Testing Commission (PTC)

Contents

* Corresponding author. Tel.: +45 35 32 61 13; fax: +45 35 32 62 70.
E-mail address: niels.morling@forensic.ku.dk (N. Morling).

1872-4973/$ – see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2007.06.006

Please cite this article in press as: D.W. Gjertson et al., ISFG: Recommendations on biostatistics in paternity testing, Forensic Science International: Genetics (2007), doi:10.1016/j.fsigen.2007.06.006

1. Introduction

In 2002, the Paternity Testing Commission (PTC) of the International Society for Forensic Genetics (ISFG) published a set of recommendations based on ISO 17025 standards concerning selected areas of importance to paternity testing [1]. The ISO 17025 standards are stated in general terms, but their guidelines anticipate the need for attaching explanations (AKA applications) in specialty areas so long as explanations do not include additional requirements [2]. The PTC identified areas that needed explanations and added these explanations as recommendations related to paternity testing, conforming to wording of previous publications of recommendations of the ISFG. Although the recommendations clarified ISO standards as they applied to paternity testing, they did not specifically focus on the biostatistical evaluation of paternity testing except to add the following limited explanation to Section 5.10.2 'Test reports and calibration certificates':

If the weight of the evidence is calculated, it shall be based on likelihood ratio principles.

The paternity index (PI) is a likelihood ratio:

$$\mathrm{PI} = \frac{\text{probability(types observed}|\text{the hypothesis is that the tested man is the father)}}{\text{probability(types observed}|\text{the hypothesis is that a random man is the father)}}$$

If other values on likelihood ratio principle are presented, e.g., Wahrscheinlichkeit W, the premises and assumptions shall be clearly specified.

For completeness, the denominator of PI may also be stated as the probability (types observed |the hypothesis that the tested man is unrelated to the father).

In 2004, the board of the ISFG appointed the Paternity Testing Commission of the ISFG to establish specific recommendations on biostatistics in paternity testing. The purpose of this report is to provide practical explanations regarding the implementation of likelihood ratio principles to summarize the genetic evidence in paternity testing. To help set up the recommendations, we begin with a brief review on the biostatistical evaluation of disputed parentage cases.

2. Abbreviated historical background

In much the same manner as Mendel's work was originally ignored, Essen-Möller's investigations [3,4] into positive proof of paternity went largely unrecognized for nearly 20 years. Essen-Möller and his mathematical colleague Quensel devised a formula (generally known as the Essen-Möller formula) for standard paternity cases involving a putative father, mother and child, which enabled serological phenotypes to be expressed numerically as a probability of paternity. They arrived at the relation $W = X/(X + Y)$ in which the terms X and Y represented probabilities of the hypotheses "paternity" and "non-paternity," respectively. Applying this formulation to a defined population, one identifies constellations composed of mother–child pairs with the same phenotypes as the mother and child. There will be a fraction, X, of these constellations that have the true father with the same phenotypes as the putative father. Also, in the overall population, there will be a fraction, Y, of men with identical phenotypes to the putative father. W gives the probability of the putative father belonging to the subset of fathers. Box 1 illustrates the logic of their approach.

Twenty-three years after Essen-Möller published his formula, Ihm showed that the formula could be derived from a straightforward application of Bayes' theorem using the Bayes' postulate of equal a priori probabilities for and against paternity [5]. In 1956, Gürtler proposed the ratio $\mathrm{PI} = X/Y$ as a basic index to report the likelihood of paternity with large values suggesting fatherhood [6]. Later, Ihm showed that the best test of the "non-paternity" hypothesis (H_0) has the form: reject H_0 if $\mathrm{PI} > (\pi_0 A)/(\pi_1 B)$ provided $\pi_1 > 0$, where π_0 and π_1 are the a priori probabilities against and for paternity, respectively [7]. When $\pi_0 = \pi_1$, the test is conducted by setting A/B to be consistent with laboratory criteria and local policies. Typical values of A/B are 100 or 1000. Thus, if PI is greater than A/B based on an initial battery of tests, a case report is issued; otherwise, further genetic tests are executed until either the criteria are reached or all possible tests are exhausted. Valentin observed that PI is a sufficient statistic since a case's distribution of observed phenotypes (given PI) does not depend on whether the putative father is the true father that is, it embodies all the information that is available from knowledge of the genotypes [8].

Alternative methods for deciding paternity have been reported. Wiener [9] proposed a probability of paternity which differs from the Essen-Möller version. Wiener's formula is based on exclusion probabilities derived from the phenotypes of

Box 1. Suppose the putative father (PF) is AB, the mother (MO) is CD and the child (CH) is BC for a co-dominant system of alleles A, B, C, D and E, where E is a marker representing the collection of all other markers not in the set {A,B,C,D}. Let A, B, C, D's frequencies equal a, b, c, d, respectively and let E's frequency $e = 1 - (a + b + c + d)$. Essen-Möller postulated that X is the fraction of constellations that have the true father with the same phenotypes as PF, and, for our example, this would be given by:

$$X = \frac{2ab \cdot 2cd \cdot 1/2 \cdot 1/2}{2ab \cdot 2cd \cdot 1/2 \cdot 1/2 + b^2 \cdot 2cd \cdot 1 \cdot 1/2 + 2cb \cdot 2cd \cdot 1/2 \cdot 1/2 + 2bd \cdot 2cd \cdot 1/2 \cdot 1/2 + 2be \cdot 2cd \cdot 1/2 \cdot 1/2} = \frac{abcd}{bcd} = a.$$

Next, Y is the fraction of men with identical phenotypes to PF and equals:

$$Y = \frac{2ab}{1} = 2ab.$$

Thus, Essen-Möller's paternity probability equals

$$W = \frac{X}{(X+Y)} = \frac{a}{a+2ab} = \frac{1}{1+2b}.$$

If $b = 0.2$, say, then, $W \approx 71\%$.

Please cite this article in press as: D.W. Gjertson et al., ISFG: Recommendations on biostatistics in paternity testing, Forensic Science International: Genetics (2007), doi:10.1016/j.fsigen.2007.06.006

the mother–child pair and does not make direct use of the putative father's phenotypes. This is one of several respects in which an ''exclusion'' approach is inferior. Various other derivations for a probability of paternity have been proposed including a ''Neyman–Pearson'' type method by Schulte-Mönting and Walter [10] and a minimax method by Ihm [7].

In 1985, Li and Chakravarti claimed that the Essen-Möller Probability of Paternity was invalid and that PI was not a likelihood ratio [11]. However, their claims were refuted by a series of papers that validated PI's formulation [12–14]. Elston demonstrated that Essen-Möller's version was more efficient (smaller mean squared error) than Wiener's method simply because W makes use of more information [12], and Baur et al. proved that PI is a proper likelihood ratio – the ratio of two probabilities of an observable event (the phenotypic constellations of the tested individuals) conditional on two mutually exclusive hypotheses [13]. In the last paper of this series, Mickey et al. demonstrated that W and PI were valid measures of paternity based on actual casework [14].

Generally, PI requires tedious calculation. However, two rules – Hardy–Weinberg equilibrium (HWE) and the product rule – ease computation allowing efficient schemes and algorithms to be derived. Early on, Hummel, in cooperation with Ihm and Schmidt, tabled W and log(PI] values for all possible putative father–mother–child phenotypic constellations in low-polymorphic genetic systems [15]. Building on work by Wehner and Rittner [16], Baur et al. [17] produced computer programs that accommodated high-polymorphic systems as well as a wide variety of parentage situations in addition to the ''standard'' paternity case. Brenner extended these ideas to produce a computer program that calculates likelihoods of alleged genetic relationships among any miscellaneous collection of people [18].

The product rule allows PI values derived from independent genetic systems to be combined into a total likelihood ratio (PI_N) by simple multiplication. Symbolically, the combined paternity index equals

$$PI_N = \prod_{i=1}^{N} PI_i$$

where $i = 1, \ldots, N$ indexes the individual genetic systems. Caution should to be exercised when implementing the product rule since genetic systems may be dependent because (1) they are physically linked when expressed on the same chromosome (linkage disequilibria among loci[1]) or (2) their genes may exhibit non-random assortment even when expressed on different chromosomes due to population substructure. Although methods exist for incorporating linkage disequilibria (e.g., HLA haplotype analysis) and population substructure (e.g., inbreeding coefficients) into genetic calculations [19,20], a large body of evidence indicates that, for the major racial/ethnic groups and for the routine markers used in parentage testing, corrections do not significantly alter PI's accuracy [21–25].

[1] Physical linkage implies linkage within a family, but while this may be a concern for some kinship problems it is not generally relevant for paternity. Thanks to the opportunity for recombination over many generations, linkage disequilibrium within a population – the possible concern for paternity – is not observed in major populations unless the systems are extremely close, such as the various HLA loci.

3. Proposed ISFG-recommendations

The ISFG recommendations on biostatistics in paternity testing are organized as follows.

R1 Mathematics
- R1.1 LR principle
- R1.2 Mutually exclusive hypotheses
- R1.3 Calculation concerns
  - R1.3.1 Possible mutation
  - R1.3.2 Possible null allele

R2 Population genetics
- R2.1 Allele probabilities
- R2.2 Y-chromosome
- R2.3 Mitochondrial DNA
- R2.4 Population substructure

R3 Special cases
- R3.1 Deficiency/reconstruction
- R3.2 Immigration

R4 Non-paternity

R5 Documentation
- R5.1 Test Reports
- R5.2 Assumptions

The specific recommendations and guidance are:

R1 Mathematics

R1.1 Likelihood ratio

The weight of the evidence shall be calculated based on likelihood ratio principles.

R1.2 Mutually exclusive hypotheses

The biostatistical evaluation of parentage shall be founded on mutually exclusive hypotheses regarding the parentage of the child or the disputed genetic relationship.

Guidance: The hypotheses are mutually exclusive, limited statements regarding the ''cause of the child'' or ''the biological relationship that exists among tested individuals.'' Hypotheses represent different pedigrees that depict genetic results for individuals linked by given (for example, the mother–child) and postulated (for example, the tested man–child) relationships. Likelihood ratios are defined by contrasting hypotheses. Examples of contrasting hypotheses include ''the tested man is the father of the child versus an unrelated untested man is the father of the child'' and ''two tested individuals are siblings versus they are half-siblings.'' From each pedigree, a likelihood or conditional probability of observing the linked test results is computed based on Mendelian principles and underlying gene/haplotype frequencies. Finally, likelihoods are compared.

When the number of postulated pedigrees is limited to two, one computes a single ratio comparing the main hypothesis to

Please cite this article in press as: D.W. Gjertson et al., ISFG: Recommendations on biostatistics in paternity testing, Forensic Science International: Genetics (2007), doi:10.1016/j.fsigen.2007.06.006

the alternative, and this likelihood ratio or paternity index (PI) corresponds to one single posterior probability W

$$W = \frac{\mathrm{PI}(\pi_1/\pi_0)}{[1 + PI(\pi_1/\pi_0)]}$$

where (π_1/π_0) represents the prior odds in favor of the main hypothesis.

The number of well-defined hypotheses need not be limited to just two. Three or more hypotheses may be naturally generated. For example, sometimes these mutually exclusive hypotheses are of interest: (1) the putative father is the father of the child; (2) the putative father's brother is the father of the child; (3) the putative father is unrelated to the father of any child. If there are n hypotheses, then there are n likelihoods which may be considered pairwise as $n * (n - 1)/2$ likelihood ratios. As usual, interpreting these requires consideration of prior probabilities. Let L_i represent the likelihood of the ith pedigree, $i = 1, \ldots, n$. Then,

$$W_i = \frac{\pi_i L_i}{\sum_{j=1}^{n} \pi_j L_j}$$

where π_j represents the prior probability of jth pedigree and $\sum \pi_j = 1$.

From the legal perspective, genetic tests corroborate an accusation of paternity or other genetic relationship that originates from external evidence. Practically, however, laboratories do occasionally formulate hypotheses *post hoc*. When *post hoc* hypotheses are considered, the laboratory shall be cautious about suggesting prior probabilities and cognizant of the possibility of experimenter bias. The example shown in Box 2 helps illustrate the interpretation of multiple hypotheses described above. [Also see Brenner [26] for guidance involving hypotheses in disaster victim identification situations and other complex kinship cases.]

R1.3 Calculation concerns [regarding special circumstances in standard and non-standard cases requiring additional consideration in the biostatistical evaluation of paternity testing)

R1.3.1. Possible mutation

PI shall be modified for possible mutation patterns between tested individuals when isolated mismatches among tested systems, which normally lead to an opinion of non-relationship, may be the result of mutant DNA causing false conclusions. The method for modifying PI shall be documented.

Guidance: The possibility of mutation shall be taken into account whenever a genetic inconsistency is observed. For typical DNA systems, average mutation rates range from 0.005 to 1% [27]. Also, mutation probabilities are affected by sex as well as DNA fragment sizes [27,28].

The theory is illustrated using the following hypothetical case: putative father's, mother's and child's alleles equal (A,B), (C,D) and (C,E), respectively. Let $\mu_{I,J}$ equal the specific mutation rate for changing allele I to J where I and J come from the set of DNA markers. Given the specific transition probabilities and ignoring multiple mutational events, $X = (1/2)(1/2)\ (\mu_{\mathrm{A,E}} + \mu_{\mathrm{B,E}})$ and $Y = (1/2)e$, where e represents the probability for the "E" allele in a defined population. Thus,

$$\mathrm{PI} = \frac{X}{Y} = \frac{(\mu_{\mathrm{A,E}} + \mu_{\mathrm{B,E}})}{2e}$$

Unfortunately, current estimates for $\mu_{I,J}$ are equivocal for most DNA systems.

Several investigators have offered practical solutions involving the use of average mutation rates per system (μ_{m}), which avoid the use of specific mutation rates [29–31]. For example, Brenner suggested adjusting PI based on μ_{m} and a probability distribution for the number of repeat units altered during the mutational event [31]. His method is especially well suited for STR systems.

Due to the presence of fragment-band measurement error, small mutations in RFLP loci may be indistinguishable from gel resolution limitations. Thus, RFLP systems are especially problematic with regard to specific mutation rates. Gjertson [32] and others suggested substituting the following average formula for RFLP PI, which is a modified version of one published by Fimmers et al. [30]:

$$\mathrm{PI}_{\mathrm{RFLP}} = \frac{\mu_{\mathrm{m}}}{A}$$

where A equals the average probability of exclusion for non-fathers in the given system. In practice, average mutation PI shall be calculated by each laboratory using its empirically validated data since values can vary depending on experimental conditions.

Box 2. Example: A cumulative PI = 94 is obtained from 15 tested loci: PI = 255,000 at the 14 "consistent" loci, and PI = 1/2700 from vWA where Child = 16 and Tested man = 14, 18 so a 2-step mutation is necessary to explain paternity. We decide to consider the third possibility of unclehood as mentioned above. The likelihood ratios favoring the hypotheses (1) the putative father is the father of the child, (2) the putative father's brother is the father of the child, (3) the putative father is unrelated to the father of any child, respectively over (3) are 94 (the "PI"), 700 (the avuncular index or "AI"), and of course 1. Then these three numbers may be considered the likelihoods L_i, $i = 1, 2, 3$.
Prior probabilities: For the sake of illustration assume (see §R5.2) prior probabilities π_i of 50, 2, and 48%.
Calculation:

Hypothesis	Father	Uncle	Unrelated
Likelihoods, L_i	94	700	1
Priors, π_{I}	50%	2%	48%
Relative posteriors, $\pi_i L_i$	47	14	0.48
Posterior probabilities $\pi_i L_i / \Sigma \pi_i L_i$	76%	23%	1%

Note that 76% compares to a posterior probability of nearly 99% if the "uncle" hypothesis is not considered. The conclusion is quite sensitive to the prior probability π_2 hypothesized for the "uncle" possibility.

R1.3.2. Possible null allele

The probability of silent and null alleles shall be considered in biostatistical calculations.

Guidance: If only a single allele or single apparent haplotype is identified in a tested subject, the possibility of a silent or null allele and/or haplotype shall be considered when calculating PI. A null allele is an allele that does not contribute to the phenotypic result of the test. STR null alleles mainly occur when a DNA-primer in the PCR reaction fails to hybridize to the template DNA and thereby results in no detectable PCR product. Laboratories should try to resolve this situation with different pairs of STR primers.

Persistent null alleles are rare for standard STR systems because of good primer design, and therefore accurate null allele frequencies are elusive. If a null allele is critical for evaluation of a particular case – namely if child and alleged father are differently homozygous – and no data from which to estimate the null allele frequency exists, consider a generous bracket of plausible values for the frequency and compute a corresponding range of values for the PI. If the resulting range of case-specific PI_N values are all large (as large as a laboratory's threshold value for issuing non-exclusion paternity reports), then report that the PI is "at least greater than the smallest value" in the range, or, if the resulting range of PI_N values are all small (smaller than the threshold value), then report that the PI is "at most less than the largest value" in the range. For traditional systems such as the HLA and ABO systems, maximum likelihood estimates of null alleles can be made using the EM algorithm, which is analogous to iterative gene-counting [33], by constructing databases of independent observed phenotypes rather than inferred genotypes containing apparent homozygotes.

R2 Population genetics

R2.1 Allele probabilities

The probability of observing an allele, i, can be estimated as:

$$\frac{x_i + 1}{N + 1}$$

where x_i is the number of i alleles and N is the total number of alleles in the existing database.

Guidance: The relevant probability of observing an allele is its conditional probability given observation among tested individuals. The database sample frequency of x_i/N, ignoring a new observation in a tested trio, is regularly biased toward paternity.[2] Extending the database with one extra observation is a simple and nearly accurate procedure to overcome the bias. In particular, occasionally, a new allele not present in a reference database is observed in routine testing. Then the formula reduces to $1/(N + 1)$ since the marker went unobserved among the N previous alleles in the database. Additionally, laboratories may choose to follow a minimum count policy, such as the NRC II recommendation of a minimum numerator of 5 [34].

R2.2 Y-chromosome

Results from Y-chromosome markers shall be handled as haplotypes, and haplotype probability estimates shall be used for calculation.

Guidance: Likelihoods for the Y-chromosome must be based on haplotype frequencies. See Gill et al. [35] and Gusmao et al. [36] for recommendations regarding Y-chromosome reference databases. The weight of the genetic evidence of Y-chromosome markers shall be combined with the genetic weight from independent, autosomal genetic markers. Since Y-chromosome markers in males are similar from generation to generation except for mutation events, they are generally informative only in cases where no other family members in the paternal lineage are relevant for alternative hypotheses. In cases where family members are relevant for alternative hypotheses, LRs from Y-STRs or Y-SNPs are often close to one and, therefore, are useless as genetic evidence. [For completeness, the biostatistical evaluation of X-chromosome markers is purposely not addressed since laboratories are not routinely using X-linked markers in relationship testing. See Szibor et al. [37,38] for recent discussions of the use of X-chromosome markers.]

R2.3 Mitochondrial DNA

Results from mitochondrial DNA (mtDNA) markers shall be handled as single entities per subject, and pseudo-haplotype probabilities shall be used for calculation.

Guidance: MtDNA markers are passed en bloc from generation to generation by maternal inheritance and, thus, behave as apparent haplotypes-defined here as pseudo-haplotypes. MtDNA is similar from generation to generation except for mutation events. Therefore, a PI for mtDNA markers must be based on pseudo-haplotype frequencies. See Carracedo et al. [39] for guidelines regarding the construction of mtDNA reference databases. Please note that there are as yet no consensual approaches to estimate population-specific mtDNA frequencies [40,41]. The weight of the genetic evidence of mtDNA markers shall combined with the genetic weight from independent, autosomal genetic markers by multiplication of contributions to the likelihood ratio. Independence is key here, and caution is warranted in simple multiplication based on the grounds of population substructure [see Buckleton et al. [42]]. In addition, only in cases where no other family members in the mtDNA lineage are relevant for alternative hypotheses can the genetic evidence of mtDNA markers add substantially to the overall results from autosomal genetic markers.

R2.4 Population substructure

If a significant degree of substructuring is known to be present in a population, algorithms that take substructuring into consideration shall be used.

Guidance: General algorithms for incorporating population

[2] In recessive and linked-loci systems, the bias may lie either for or against paternity.

substructure into PI have been presented by Evett and Weir [20]. [Also see Gjertson and Morris [43] for presentation of an alternative to the product rule for PI that offers a means of avoiding objections to the rule.] As stated in Background, in many populations the degree of substructuring is so small that, for practical purposes, it does not affect the biostatistical evidence to any significant degree in paternity testing. If no significant substructuring exists in a population, the biostatistical calculations can be performed without correcting for substructuring [23].

R3 Special cases

R3.1 Deficiency/reconstruction

The basic principles for the biostatistical calculations in uncomplicated (standard) and complicated (deficiency or reconstruction) cases are the same.

Guidance: All biostatistical calculations in paternity testing shall be based on a likelihood ratio principle requiring the evaluation of two relevant, mutually exclusive hypotheses. This principle remains the same regardless of the complexity of the paternity case and whether or not pertinent individuals are tested. Deficiency cases include, but are not limited to, ones in which the mother is not tested, or the putative father is not tested and his genetic markers are deduced from those of relatives. The common themes among these cases are that genetic information is missing, the missing information requires statistical imputation, and imputation generally increases the complexity of formulas for *X* and *Y*. For example, one needs to consider two possible paternal alleles when the mother is not tested, whereas usually one paternal marker is central to standard cases. Valentin [44] and Asano et al. [45] formulated the general procedures in motherless cases, and Ihm and Hummel [46] and Asano et al. [47] outlined methods for PI using markers from relatives of deceased putative fathers. Comprehensive reviews for calculating likelihood ratios for almost any type of dispute have been prepared by Baur [48] and Brenner [18].

R3.2 Immigration

All biostatistical calculations in immigration cases shall be based on likelihood ratio principles.

Guidance: In all matters of relationship testing, including immigration cases, mutually exclusive hypotheses must be clearly defined. Based on these hypotheses, likelihoods of the genetic evidence are calculated, and the ratio of any two likelihoods is computed. The basic principles for the biostatistical calculations in e.g., immigration cases are similar to those in paternity testing.

In immigration cases, the clients' questions are often more complex than in paternity cases. In a paternity case, the great majority of problems can be reduced to the question whether the investigated man or an alternative man is the father of the child. In immigration cases, the prior probability of a tested individual being a close relative to a true father is often significant. When an adult has been excluded as the father or mother of a child, the client often wants to know if another close relationship exists. Therefore, it may be tempting to perform biostatistical calculations under a number of hypothetical relationships, e.g., the investigated man is the true father's brother, half-brother, father, etc, to find the explanation associated with the highest likelihood ratio. The commission advocates to exercise restraint in testing speculative hypotheses without specific knowledge about the case's circumstances, as such hypotheses may be unfounded and could incriminate the tested individuals. Also, the commission recommends that laboratories be particularly cautious about the assignment of prior probabilities and cognizant of the possibility of experimenter bias when assessing post hoc hypotheses.

R4 Non-paternity: Considerations on strategies regarding genetic evidence against paternity.

Laboratories are responsible for establishing and admitting their exclusion criteria. Preference shall be given to criteria stated in terms of a PI threshold.

Guidance. These biostatistical recommendations direct laboratories to always calculate a combined PI, which will be greater than zero regardless of the number of observed genetic consistencies or inconsistencies when possible mutations, null alleles and measurement errors are considered. This raises the question of whether laboratories should round PI to zero at some point based on very small values? The ideal answer is ''no'' since, by symmetry, we do not advocate rounding PI to infinity or *W* to 100% either. Practically, however, deterministic statements of non-paternity are legal norms, and, it is easily deemed foolish to require a laboratory to report $PI > 0$ when, say, 13/13 STR systems are inconsistent. Thus, laboratories are responsible for establishing and admitting their exclusion criteria. Preference should be given to policy stated in terms of a PI threshold (e.g., $PI < 1/1000$), but countenance is given to ones based on some number of inconsistencies. Regarding the latter, a policy to exclude on e.g., three inconsistencies out of 13 CODIS STR loci means on average excluding with $PI = 1/4600$; two inconsistencies corresponds to $PI = 1/4.7$ [49].

R5 Documentation

R5.1 Test reports

In addition to the combined PI, test reports shall also contain the individual PI's for each genetic system reported and the racial/ethnic backgrounds used by the laboratory for calculations. If the probability of paternity (*W*) is reported, then the prior probability assumption used to calculate *W* shall be stated. Test reports shall include statements of assumptions, validation and computational techniques whenever alternative biostatistical methods to PI are used.

Guidance: Biostatistical results shall be presented in sufficient detail to facilitate recalculation. The reporting of individual PI's allows examination of calculation principles and review of concordant results when two laboratories use different sets of genetic markers. Information on racial/ethnic backgrounds helps to delineate the reference databases.

Please cite this article in press as: D.W. Gjertson et al., ISFG: Recommendations on biostatistics in paternity testing, Forensic Science International: Genetics (2007), doi:10.1016/j.fsigen.2007.06.006

The value of the scientific evidence is the likelihood ratio (PI). In addition to reporting this number, a laboratory may include posterior probabilities (W) calculated at various prior probabilities (π_1) as examples. However, the laboratory shall not itself assume *the* prior probability. The following summary statement is acceptable: "PI = 200 and $W = 99.5\%$ assuming 50% probability *a priori*." While monitoring the distribution of PI can justify certain prior probabilities for empirical validation studies [14], such a "laboratory prior" is not properly taken as a stand-in for the facts of a particular case. [See Potthoff and Whittinghill [50], Hummel et al. [51], Baur et al. [52] for additional arguments in the assignment of a priori probabilities.] Laboratories may present W at a variety of different prior probabilities in test reports, and any prior probabilities of paternity shall be stated along with W to allow assessment of prior weights assigned to the mutually exclusive hypotheses.

Confidence intervals of PI and W are irrelevant in probability estimates and hence unnecessary. Also, 'conservative' estimates of the weight of genetic evidence are generally irrelevant to paternity testing as it is unfounded (from the laboratory's point of view) to assert a conservative side. In contradistinction and as mentioned below, it may be appropriate to test W's sensitivity to changeable assumptions (for example, altering prior probabilities for and against paternity) and present ranges of posterior probabilities to aid in the interpretation of the biostatistical evidence.

R5.2 Assumptions

> Laboratory procedures shall document assumptions and validate frequencies used to compute PI. The reference database shall be selected so that it can be used for estimation of the probability of obtaining the genetic results under the assumption of the relevant hypotheses. If a threshold PI exists for issuing test reports, the value(s) shall be documented.

Guidance: All mathematical calculations require assumptions. The laboratory shall document PI*s* assumptions to ensure quality and client satisfaction. Several assumptions are necessary to compute and interpret PI, and they can be categorized as fundamental, empirical, specific and changeable.

Fundamental claims imply correctness in laws of genetics and mathematics. These laws are derived from basic principles and are usually accepted as true or appropriate without justification.

Gene and haplotype probabilities are estimated through *empirical* sampling of populations. The laboratory shall validate and document that their frequency databases are representative of defined populations (for a review of scientific standards in forensic genetics, see Ref. [53]). Validation evidence includes, but is not limited to, the following examples.

- When utilizing published frequency tables, concordance between frequencies estimated from randomly sampled persons typed by the laboratory and the tabled frequencies shall be demonstrated.
- When importing databases from another laboratory, concordance in pertinent test results of a random sample of a sufficient number of individuals shared between laboratories shall be demonstrated. This may be achieved by participating in proficiency testing exercises covering the genetic systems in question.

Specific assumptions are those made to limit relationships among tested individuals and qualify test results in order to produce specific formulas for PI. In paternity cases, it is usually assumed that subjects have been accurately identified, maternity is undisputed, mating is random (where possible fathers are not related to the mother or to each other), and phenotypes have been accurately determined without error. Furthermore, the individually reported genetic systems are assumed independent so that the product rule applies. The exact formulas for calculating PI depend on the specific assumptions and the phenotypic constellations of the tested parties. For DNA-based tests with unambiguous genotype assignment, tables of formulas exist for possible allele-sharing patterns, greatly simplifying the calculation [54]. If measurement error exists, then the error shall be incorporated in the formulas for PI greatly increasing their complexity [55,56]. Also, formulas have been developed to handle exceptional cases, such as incest and possible fathers who are related to each other [57]. In incest cases involving tested subjects, formulas for paternity indices change only when the incestuous relationship between the putative father and mother yields information regarding their genotypes (e.g., the phase of HLA haplotypes, the existence of recessive alleles). In the absence of a putative father's sample, formulas for incest indices assess the possibility of an incestuous relationship producing a child (as opposed to a particular man fathering the child).

Race and prior probabilities constitute the *changeable* assumptions.[3] Race or ethnic group is intended to define a population in the formulation of PI, and gene/haplotype frequencies may vary markedly from one group to the next. In practice, a subject's race/ethnicity is assigned by interview, and the alternative father's race/ethnicity is usually equated with that of the putative father so frequencies are tallied for that group. However, the alternative hypothesis may include an assumption of the population to which the alternative father belongs. Thus, the tested man and the alternative man may belong to two different populations. In practice, this situation usually arises in cases of mixed races or ethnicities between mother and putative father. Justifying one race/ethnicity may be difficult, and, in such cases, it is appropriate to test and report PI*s* sensitivity by calculation under a variety of assumptions.

[3] Here, changeable assumptions are regarded as distinct from specific assumptions. The distinction reflects the type of information available to support an assumption. Specific assumptions are usually fixed by objective, tangible evidence, whereas changeable assumptions may vary due to subjective information. For example, medical records can support "undisputed maternity," but personal beliefs underlie "prior probabilities." Changeable assumptions like the calculation race and prior should be noted on a test report, whereas specific assumptions are usually implicit and not noted.

Please cite this article in press as: D.W. Gjertson et al., ISFG: Recommendations on biostatistics in paternity testing, Forensic Science International: Genetics (2007), doi:10.1016/j.fsigen.2007.06.006

Acknowledgements

We thank Charlotte Hallenberg, MSc, PhD for helpful discussions and assistance with the organization of the work in the commission. We thank Professor, Dr.med. Christian Rittner for hosting the meeting of the Paternity Testing Commission and for helpful discussions. We thank Dr. Toshimichi Yamamoto, who represented the Japanese Speaking Working Group at the meeting of the Paternity Testing Commission, for helpful discussions.

References

[1] N. Morling, R. Allen, A. Carracedo, H. Geada, F. Guidet, C. Hallenberg, W. Martin, W. Mayr, B. Olaisen, V. Pascali, P.M. Schneider, Paternity Testing Commission of the International Society of Forensic Genetics: recommendations on genetic investigations in paternity cases, Forensic Sci. Int. 129 (2002) 148–157.

[2] EN ISO/IEC 17025: 1999 Standard 'General requirements for the Competence of Testing and Calibration Laboratories,' International Standardization Organization, Geneva, Switzerland, 1999.

[3] E. Essen-Möller, Die Beweiskraft der Ähnlichkeit im Vaterschaftsnachweis - theoretische Grundlagen, Mitt. Anthrop. Ges. (Wien) 68 (1938) 9–53.

[4] E. Essen-Möller, C. Quensel, Zur Theorie des Vaterschaftsnachweises aufgrund von Ähnlichkeitsbefunden, Z. Ges. Gerichtl. Med. 31 (1939) 70–96.

[5] P. Ihm, Die mathematischen Grundlagen, vor allem für die statistische Auswertung des serologischen und anthropologischen Gutachtens, in: K. Hummel (Ed.), Die medizinische Vaterschaftsbegutachtung mit biostatistischem Beweis, Fischer, Stuttgart, 1961, pp. 128–145.

[6] H. Gürtler, Principles of blood group statistical evaluation of paternity cases at the University Institute of Forensic Medicine Copenhagen, Acta. Med. Leg. Soc. (Liege) 9 (1956) 83–94.

[7] P. Ihm, The problem of paternity in the light of decision theory, in: K. Hummel, J. Gerchow (Eds.), Biomathematical Evidence of Paternity, Springer-Verlag, Berlin, Heidelberg, 1981, pp. 53–68.

[8] J. Valentin, Positive evidence of paternity calculated according to Essen-Möller: the Bayesian approach, in: R. Walker (Ed.), Inclusion Probabilities in Parentage Testing, American Association of Blood Banks, Arlington, Virginia, 1983, pp. 63–75.

[9] A. Wiener, Likelihood of parentage, in: L. Sussman (Ed.), Paternity Testing by Blood Grouping, Charles C. Thomas, Springfield, Illinois, 1976, pp. 124–131.

[10] J. Schulte-Mönting, E. Walter, Statistische Interpretation von serologischen Befunden zur Beurteilung einer Vaterschaft, Bundesgesundheitsbl. 18 (1972) 257–259.

[11] C. Li, A. Chakravarti, Basic fallacies in the formulation of the paternity index, Am. J. Hum. Genet. 37 (1985) 807–818.

[12] R. Elston, Probability and paternity testing, Am. J. Hum. Genet. 39 (1986) 112–122.

[13] M. Baur, R. Elston, H. Gürtler, K. Henningsen, K. Hummel, H. Matsumoto, W. Mayr, J. Morris, L. Niejenhuis, H. Polesky, D. Salmon, J. Valentin, R. Walker, No fallacies in the formulation of the paternity index, Am. J. Hum. Genet. 39 (1986) 528–536.

[14] M. Mickey, D. Gjertson, P. Terasaki, Empirical validation of the Essen-Möller probability of paternity, Am. J. Hum. Genet. 39 (1986) 123–132.

[15] K. Hummel (Ed.), Biostatistical Opinion of Parentage, Gustav Fischer Verlag, Stuttgart, 1971.

[16] H. Wehner, C. Rittner, Das Bayessche Theorem und die Überprüfung seiner Anwendung zur Berechnung der Vaterschaftsplausibilität, Z. Rechtsmedizin 69 (1971) 125–131.

[17] M. Baur, W. Mayr, C. Rittner, Algorithm for the computation of plausibilities of paternity in the HLA system, Z. Immunitaetsforsch 152 (1976) 209–219.

[18] C. Brenner, Symbolic kinship program, Genetics 145 (1997) 535–542 (Published erratum appears in Genetics 147 (1997) following 398.).

[19] D. Balding, R. Nichols, A method for quantifying differentiation between populations at multi-loci and its implications for investigating identity and paternity, Genetica 96 (1995) 3–12.

[20] I. Evett, B. Weir, Interpreting DNA Evidence. Statistical Genetics for Forensic Scientists, Sinauer Associates, Sunderland, Massachusetts, 1998

[21] B. Devlin, N. Risch, K. Roeder, No excess of homozygosity at loci used for DNA fingerprinting, Science 249 (1990) 1416–1420.

[22] R. Chakraborty, K. Kidd, The utility of DNA typing in forensic work, Science 254 (1991) 1735–1739.

[23] N. Morton, Genetic structure of forensic populations, Proc. Natl. Acad. Sci. USA 89 (1992) 2556–2560.

[24] J. Morris, D. Gjertson, The paternity index, population heterogeneity, and the product rule, in: W. Bar, A. Fiori, U. Rossi (Eds.), Advances in Forensic Haemogenetics, vol. 5, Springer-Verlag, Berlin, 1993, pp. 435–437.

[25] D. Stivers, R. Chakraborty, A test of allelic independence based on distributions of allele size differences at microsatellite loci, Hum. Hered. 47 (1997) 66–75.

[26] C. Brenner, Some mathematical problems in the DNA identification of victims in the 2004 tsunami and similar mass fatalities, Forensic Sci. Int. 157 (2006) 172–180.

[27] B. Brinkmann, M. Klintschar, F. Neuhuber, J. Hoehne, R. Burkhard, Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat, Am. J. Hum. Genet. 62 (1998) 1408–1415.

[28] Appendix 11: Mutation rates summarized for genetic markers analyzed by RFLP mapping and PCR. p. 146–148 Guidance for Standards for Parentage Testing Laboratories, seventh Edition, AABB, Arlington, Virginia, 2006, pp. 146–148.

[29] A.P. Dawid, J. Mortera, V.L. Pascali, Non-fatherhood or mutation? A probabilistic approach to parental exclusion in paternity testing, Forensic Sci. Int. 124 (2001) 55–61.

[30] R. Fimmers, L. Henke, J. Henke, M. Baur, How to deal with mutations in DNA-testing, in: C. Rittner, P.M. Schneider (Eds.), Advances in Forensic Haemogenetics, vol. 4, Springer-Verlag, Berlin, 1992, pp. 285–287.

[31] C. Brenner, Mutations in paternity, 2006, http://dna-view.com/mudisc.htm.

[32] D. Gjertson, Appendix 12: The effect of isolated single-locus inconsistencies in the statistical evaluation of paternity: A 2005 update. Guidance for Standards for Parentage Testing Laboratories, seventh ed., AABB, Arlington, Virginia, 2006, pp. 152–157.

[33] R. Ceppellini, M. Siniscalco, C. Smith, The estimation of gene frequencies in a random-mating population, Ann. Hum. Genet. 20 (1955) 97–115.

[34] National Research Council, The evaluation of forensic DNA evidence/ Committee on DNA Forensic Science: an Update, Commission on DNA Forensic Science: an Update, National Academy Press, Washington, DC, 1996.

[35] P. Gill, C. Brenner, B. Brinkmann, B. Budowle, A. Carracedo, M.A. Jobling, P. de Knijff, M. Kayser, M. Krawczak, W.R. Mayr, N. Morling, B. Olaisen, V. Pascali, M. Prinz, L. Roewer, P.M. Schneider, A. Sajantila, C. Tyler-Smith, DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs, Int. J. Legal Med. 114 (2001) 305–309.

[36] L. Gusmao, J.M. Butler, A. Carracedo, P. Gill, M. Kayser, W.R. Mayr, N. Morling, M. Prinz, L. Roewer, C. Tyler-Smith, P.M. Schneider, International Society of Forensic Genetics. DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis, Int. J. Legal Med. 120 (2006) 191–200.

[37] R. Szibor, M. Krawczak, S. Hering, J. Edelmann, E. Kuhlisch, D. Krause, Use of X-linked markers for forensic purposes, Int J Legal Med 117 (2003) 67–74.

[38] R. Szibor, J. Edelmann, S. Hering, X-chromosomal markers for forensic purpose, 2007, http://www.chrx-str.org.

[39] A. Carracedo, W. Bär, P. Lincoln, W. Mayr, N. Morling, B. Olaisen, P. Schneider, B. Budowle, B. Brinkmann, P. Gill, M. Holland, G. Tully, M. Wilson, DNA Commission of the International Society for Forensic

Genetics: guidelines for mitochondrial DNA typing, Forensic Sci. Int. 110 (2000) 79–85.

[40] W. Parson, A. Brandstätter, A. Alonso, N. Brandt, B. Brinkmann, A. Carracedo, D. Corach, O. Froment, I. Furac, T. Grzybowski, K. Hedberg, C. Keyser-Tracqui, T. Kupiec, S. Lutz-Bonengel, B. Mevag, R. Ploski, H. Schmitter, P. Schneider, D. Syndercombe-Court, E. Sorensen, H. Thew, G. Tully, R. Scheithauer, The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives, Forensic Sci. Int. 139 (2004) 215–226.

[41] A. Salas, A. Carracedo, V. Macaulay, M. Richards, H.J. Bandelt, A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics, Biochem. Biophys. Res. Commun. 335 (2005) 891–899.

[42] J. Buckleton, C. Triggs, S. Walsh, Forensic DNA Evidence Interpretation, CRC Press, 2005.

[43] D. Gjertson, J. Morris, Assessing probability of paternity and the product rule in DNA systems, Genetica 96 (1995) 89–98.

[44] J. Valentin, Bayesian probability of paternity when mother or putative father are not tested: Formulas for manual computation, Hereditas 149 (1979) 405–416.

[45] M. Asano, K. Minakata, H. Hattori, Diagnosis of paternity for cases without the mother and without both mother and putative fathers based on blood group findings from the relatives, Z. Rechtsmed. 84 (1980) 135–144.

[46] P. Ihm, K. Hummel, A method to calculate the plausibly of paternity using blood group results of any relatives, Z. Immunitaetsforsch. 149 (1975) 405–416.

[47] M. Asano, K. Minakata, H. Hattori, General formulas of the estimated likelihood ratio Y/X in the diagnosis of paternity of a deceased putative father, Z. Rechtsmed. 84 (1980) 125–133.

[48] M. Baur, Erweiterung des Essen-Möller-Modells und die praktische Durchführung der serologisch-biostatistischen Abstammungsbegutachtung mit dem Programmsystem P.A.P.I. Bonn: Rheinische Friedrich-Wilhelms-Universität, 1977.

[49] C. Brenner, Multiple mutations, covert mutations and false exclusions in paternity casework, in: C. Doutremépuich, N. Morling (Eds.), Progress in Forensic Genetics, vol. 10, Elsevier B.V, Amsterdam, 2004, pp. 112–114.

[50] R. Potthoff, M. Whittinghill M, Maximum-likelihood estimation of the proportion of nonpaternity, Am. J. Hum. Genet. 17 (1965) 480–494.

[51] K. Hummel, O. Kundinger, A. Carl, The realistic prior probability from blood group findings for cases involving one or more men. Part II. Determining the realistic prior probability in one-man cases (forensic cases) in Freiburg, Munich, East Berlin, Austria, Switzerland, Denmark, and Sweden, in: K. Hummel, J. Gerchow (Eds.), Biomathematical Evidence of Paternity, Springer-Verlag, Berlin, 1981, pp. 81–87.

[52] M. Baur, C. Rittner, H. Wehner, The prior probability parameter in paternity testing. Its relevance and estimation by maximum likelihood. Lectures of the Ninth International Congress of the Society for Forensic Hemogenetics, Bern (1981) 389–392.

[53] P.M. Schneider, Scientific standards for studies in forensic genetics, Forensic Sci. Int. 165 (2007) 238–243.

[54] M. Traver, Appendix 8: Formulas for paternity index and RMNE values for simple codominant systems. Guidance for Standards for Relationship Testing Laboratories, seventh ed., AABB, Arlington, Virginia, 2006, p. 139.

[55] D. Gjertson, M. Mickey, J. Hopfield, T. Takenouchi, P. Terasaki, Calculation of probability of paternity using DNA sequences, Am. J. Hum. Genet. 41 (1988) 860–869.

[56] D. Berry, Inferences using DNA profiling in forensic identification and paternity cases, Stat. Sci. 6 (1991) 175–205.

[57] J. Morris, R. Garber, J. d'Autremont, C. Brenner, The avuncular index and the incest index, in: W. Mayr (Ed.), Advances in Forensic Haemogenetics, vol. 2, Springer-Verlag, Berlin, 1988, pp. 607–611.

Please cite this article in press as: D.W. Gjertson et al., ISFG: Recommendations on biostatistics in paternity testing, Forensic Science International: Genetics (2007), doi:10.1016/j.fsigen.2007.06.006