Forensic Science International: Genetics 28 (2017) 203–210

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Likelihood ratio and posterior odds in forensic genetics: Two sides of the same coin

Amke Caliebe[a,*], Susan Walsh[b,c], Fan Liu[c,d,e], Manfred Kayser[c], Michael Krawczak[a]

[a] Institute of Medical Informatics and Statistics, Kiel University, Brunswiker Strasse 10, 24105 Kiel, Germany
[b] Department of Biology, Indiana-University-Purdue-University-Indianapolis (IUPUI), 723 W. Michigan St. Indianapolis, IN, USA
[c] Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands
[d] Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, No.1 Beichen West Road, Chaoyang District, Beijing 100101, China
[e] University of Chinese Academy of Sciences, 19 A Yuquan Rd, Shijingshan District, Beijing 100049, China

ARTICLE INFO

Article history:
Received 19 August 2016
Received in revised form 2 December 2016
Accepted 4 March 2017
Available online 6 March 2017

Keywords:
Likelihood ratio
Posterior odds
Causality
Forensic DNA Phenotyping
Forensic DNA profiling
Genetic evidence

ABSTRACT

It has become widely accepted in forensics that, owing to a lack of sensible priors, the evidential value of matching DNA profiles in trace donor identification or kinship analysis is most sensibly communicated in the form of a likelihood ratio (LR). This restraint does not abate the fact that the posterior odds (PO) would be the preferred basis for returning a verdict. A completely different situation holds for Forensic DNA Phenotyping (FDP), which is aimed at predicting externally visible characteristics (EVCs) of a trace donor from DNA left behind at the crime scene. FDP is intended to provide leads to the police investigation helping them to find unknown trace donors that are unidentifiable by DNA profiling. The statistical models underlying FDP typically yield posterior odds (PO) for an individual possessing a certain EVC. This apparent discrepancy has led to confusion as to when LR or PO is the appropriate outcome of forensic DNA analysis to be communicated to the investigating authorities. We thus set out to clarify the distinction between LR and PO in the context of forensic DNA profiling and FDP from a statistical point of view. In so doing, we also addressed the influence of population affiliation on LR and PO. In contrast to the well-known population dependency of the LR in DNA profiling, the PO as obtained in FDP may be widely population-independent. The actual degree of independence, however, is a matter of (i) how much of the causality of the respective EVC is captured by the genetic markers used for FDP and (ii) by the extent to which non-genetic such as environmental causal factors of the same EVC are distributed equally throughout populations. The fact that an LR should be communicated in cases of DNA profiling whereas the PO are suitable for FDP does not conflict with theory, but rather reflects the immanent differences between these two forensic applications of DNA information.

1. Introduction

For decades, DNA profiling has served in forensic practice to facilitate the identification of trace donors. A trace DNA profile, typically comprising a selected number of highly-polymorphic short tandem repeats (STRs), is either compared to the DNA profiles of one or more suspects, or is gauged against one or more databases of DNA profiles of previously convicted persons. When a perfect match is found, i.e. when trace and target individual are of the same genotype at every STR considered, the forensic expert reports the evidential value of their result in the form of a likelihood ratio (LR).

In a forensic context, likelihoods allow weighing of the prosecution and defense hypotheses (Hp and Hd) against each other, which is not feasible by way of probabilities owing to a lack of sensible priors [1,2]. With G denoting the genetic evidence (i.e. the match between trace and target DNA profile), each likelihood is defined by the conditional probability of G given the respective hypothesis, i.e. $L(Hp|G) = P(G|Hp)$ and $L(Hd|G) = P(G|Hd)$. The likelihood ratio $LR = L(Hp|G)/L(Hd|G)$ then quantifies the relative evidential value of G with a view to decide between Hp and Hd. It

* Corresponding author.
E-mail addresses: caliebe@medinfo.uni-kiel.de (A. Caliebe), walshsus@iupui.edu (S. Walsh), f.liu@erasmusmc.nl (F. Liu), m.kayser@erasmusmc.nl (M. Kayser), krawczak@medinfo.uni-kiel.de (M. Krawczak).

http://dx.doi.org/10.1016/j.fsigen.2017.03.004
1872-4973/

must be emphasized, however, that likelihoods are not probabilities because they are not additive (i.e. the joint likelihood of some mutually exclusive hypotheses usually does not equal the sum of the individual likelihoods) and therefore fail a critical formal requirement of probability theory. Instead, the likelihoods of Hp and Hd as well as the resulting LR should be viewed as a measure of rationale belief in either of the two hypotheses.

The primary interest of the court, of course, should be in posterior odds (PO) P(Hp|G)/P(Hd|G). However, this quantity is difficult to grasp in the context of DNA profiling because of its dependence upon the prior odds P(Hp)/P(Hd) which are usually difficult to specify. Prior odds, moreover, are solely in the domain of the judge or jury. Therefore, a consensus has been reached among forensic experts that likelihoods and the LR should constitute the only case-relevant outcome of their experimental work.

In the simplest situation, a case of interest involves a single suspect and a single trace donor. Under Hd, the former would be presumed to have been drawn at random from a certain suspect population. Under Hp, the suspect is the trace donor. If the trace and suspect DNA profiles match, the LR simplifies to the inverse of the so-called '(random) match probability'. The forensic expert would then report this probability and would leave it to the court to evaluate whether the suspect left the trace or not. In principle, the same reasoning can be applied to any courtroom evidence that does not exclude a suspect or a group of suspects from trace donorship.

Forensic DNA Phenotyping (FDP) is a relatively recent development in forensic genetics. It aims at predicting selected externally visible characteristics (EVCs) of a trace donor from their DNA as left behind at the crime scene. We will continue to use the expression 'predict' in this context despite the fact that some scholars have argued that 'prediction' should only be used for future events [3]. This is because any resulting (true or perceived) logical problems can be resolved by referring to the future disclosure of the EVC of the trace donor once they have been identified. The FDP approach bears great potential in cases where DNA profiling failed, for example, because the police have no suspect at all or neither suspect DNA profile matches the trace DNA profile [4–7].

There are some major conceptual differences between DNA profiling and FDP. First, whereas identification by DNA profiling involves at least two DNA samples, namely from trace and suspect, FDP usually works with just the trace DNA. Second, FDP is not meant to yield courtroom evidence but rather to guide the police investigations in cases where DNA profiling failed or was not feasible in the first place. Most notably, from a statistical perspective, FDP yields posterior probabilities of trait phenotypes (i.e. EVCs) from genotypes by way of statistical techniques such as, for example, logistic regression analysis [8–15]. Since this seems to contradict the forensic genetics paradigm of only reporting likelihood ratios, however, some forensic experts have felt uncomfortable about reporting the PO that somebody has a certain EVC level (such as blue eye color).

Currently, practical FDP is feasible only for eye and hair color. In fact, two dedicated DNA systems have been developed and forensically validated for these EVCs, namely IrisPlex for eye color alone [12,16,17] and HIrisPlex for simultaneous eye and hair color prediction [11,18]. For skin color, efficient predictive DNA markers have been proposed as well [19], but these have not been forensically validated yet. For all other EVCs, studies to understand their genetic basis are not advanced enough to allow practical implementation of FDP [7]. This is not to say that no genetic associations have been documented yet for non-pigmentation traits such as male pattern baldness [20,21], hair structure [22] and extreme body height [23]. Rather, their prediction accuracies are not yet high enough for FDP, which would require more predictive DNA markers to be identified and added to the respective models. Moreover, for any other EVC, genetic studies have only identified the first few genes, providing very limited prediction information, or genetic studies are lacking completely (see recent overview in [7]).

In the following, we will discuss the apparent discrepancy between DNA profiling and FDP from a statistical perspective. In so doing, we will focus upon the dependency, or not, of LR and PO on population affiliation. For DNA profiling, the role of the source population of the trace donor has been well worked out before, not least including a considerable debate about the appropriateness of the so-called 'random man' assumption. Thus, it has been argued repeatedly that the population of a potential donor is rarely if ever identical to the database population used for obtaining the match probability [24,25], and a mathematical procedure known as the 'theta correction' was proposed to address this problem analytically [26–29]. So, whilst presenting a general framework of the relationship between LR and PO, we will nevertheless put our considerations into the specific context of FDP. We also propose some recommendations as to how forensic experts should report the results of their experimental work in the courtroom (DNA profiling) or to the investigating authorities (FDP).

2. Material and methods

For the present study, the performance of IrisPlex-based FDP for eye color [17] was assessed empirically in eight previously published European population samples from Norway (n = 547), the United Kingdom (n = 498), Estonia (n = 579), France (n = 616), Italy (n = 542), Greece (n = 547), Spain (n = 511) – all from the EUREYE Study [12] – and the Netherlands – from the Rotterdam Study (n = 2364) [17]. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and area under the receiver operating curve (AUC) were estimated for a prediction model obtained before from a Dutch training set [9] that did not overlap with the Dutch sample used here. All analyses were performed with the R statistics software [30], particularly packages ROCR [31] and caret [32].

3. Probabilistic models

3.1. General set-up

Couched in probability theoretical terms, every forensic application of DNA typing draws upon the causal relationship between at least two random variables, say X and Y. We will henceforth assume that X is causal for Y in the sense that the factual entity represented by X has a biological effect on that represented by Y, and that X is not merely statistically associated with Y. Various methods have been proposed in the scientific literature to arrive at a reasonable degree of belief in causality [33] and we will stipulate that, in the following, sufficient evidence for a causal relationship between X and Y exists.

In DNA profiling, X is an indicator of the identity, or not, of suspect and trace donor whereas Y is an indicator of matching DNA profiles. In FDP, X is a composite genotype whereas Y denotes the EVC of interest expressed by the unknown trace donor. Usually, composite genotype X will comprise both causal and merely associated variants so that its causality for Y is only partly warranted.

3.2. Likelihood ratio (LR) and posterior odds (PO)

Statistical inference is tantamount to the deduction from sample data of the characteristics of a so-called 'random variable', a concept borrowed from probability theory to relate chance experiments (e.g. the roll of a die or a blood pressure