

Assigning Propositions for Likelihood Ratios

*Biological Data Interpretation & Reporting Subcommittee
Biology/DNA Scientific Area Committee
Organization of Scientific Area Committees (OSAC) for Forensic Science*



OSAC Proposed Standard

Assigning Propositions for Likelihood Ratios

Prepared by
Biological Data Interpretation & Reporting Subcommittee
Version: 1.0
2018

Disclaimer:

This document has been developed by the Biological Data Interpretation & Reporting Subcommittee of the Organization of Scientific Area Committees (OSAC) for Forensic Science through a consensus process and is *proposed* for further development through a Standard Developing Organization (SDO). This document is being made available so that the forensic science community and interested parties can consider the recommendations of the OSAC pertaining to applicable forensic science practices. The document was developed with input from experts in a broad array of forensic science disciplines as well as scientific research, measurement science, statistics, law, and policy.

This document has not been published by an SDO. Its contents are subject to change during the standards development process. All interested groups or individuals are strongly encouraged to submit comments on this proposed document during the open comment period administered by the Academy Standards Board (www.asbstandardsboard.org).

Foreword

When DNA testing data are generated from evidence and then compared to data from known individuals, it is necessary to provide a statistical statement for any comparison for which an individual cannot be excluded as a contributor (or possible contributor) to the DNA. One common

$$LR = \frac{P(E|H_1, I)}{P(E|H_2, I)}$$

where:

E are the findings (e.g., DNA profile(s), presumptive test results, observations),

H_1 and H_2 are two mutually exclusive propositions, and

I is the relevant information.

The terms H_p and H_d are often used in place of H_1 and H_2 , and are assigned for the prosecution (p) and the defense (d), respectively. The arguments for the use of H_1 and H_2 in lieu of H_p and H_d , respectively, relate to avoiding any unintended assignment of propositions or scenarios to individuals such as the prosecution or defense. The arguments in favor of the alternative set, H_p and H_d , relate to directness of language. The exact propositions of the prosecution and defense may not be known, and in such circumstances, reasonable propositions consistent with both viewpoints should be selected for the analysis.

Cook et al. [5] and Evett et al. [7] classified propositions into four levels: from top to bottom these are offense, activity, source, and sub-source. This was expanded to include a sub-sub-source level and is explained in detail in Buckleton et al. [2], pp. 46-48. This document is limited to propositions at the sub-source and sub-sub-source levels.

The propositions in any given case depend on the knowledge of the case circumstances at the time the interpretation is carried out. Relevant information may include, but is not limited to, the following elements:

- a) The alleged location, direction, and time of transfer
- b) The location and time that the sample was taken
- c) The genotypes of the person of interest (POI), the complainant, and proposed consensual partners, any persons who admit activity that could deposit DNA or have legitimate access to the sample, and any elimination samples such as scene of crime staff or lab staff
- d) Evidence of opportunity for alternate donors may be relevant. In particular, the genotypes, ethnicity, or relationships¹ of persons who could reasonably be alternative donors is relevant. Also relevant is information of those who could not reasonably be considered possible donors because of lack of opportunity.

¹ There is a misperception that the likelihood ratio approach is unsuitable if any individuals within the pair of

propositions are related. This is not true as the likelihood ratio approach has been used for decades for paternity cases and missing persons. One could envisage two situations:

- a) The known individuals are related and have been genotyped, or
- b) One, or more, of the postulated unknown individuals is related to the genotyped individuals.

- e) Information of the results of other forensic analyses that suggest the presence at the scene of an alternatedonor.

When talking about the case information, it needs to be emphasized that one does not consider information such as prior conviction, motive, or a confession of the POI as relevant forensic information for the evaluation of the DNA results.

Keywords: *statistics, likelihood ratio, propositions, DNA, DNA interpretation*

DRAFT

Table of Contents

Scope.....	1
Normative References	1
Terms and Definitions.....	1
Requirements	2
Annex A (normative) Best Practice Recommendations	4
Annex B (informative) Bibliography	6

DRAFT

Assigning Propositions for Likelihood Ratios

1 Scope

This standard is to be used by laboratories for the assignment of propositions for the interpretation of DNA profiling evidence using likelihood ratios.

2 Normative References

There are no normative references.

3 Terms and Definitions

For purposes of this document, the following definitions apply.

3.1 analysis

An interpretation and/or likelihood ratio calculation.

3.2 conditioning

The act of assuming one or more pieces of information when assigning a conditional probability. The information might be the profile of an individual, or profiles of a set of individuals, that are assumed to have contributed DNA to the evidentiary item under a particular proposition, or it might simply be the assumption that a particular proposition is true. Any events (or information) that have been used for conditioning are placed to the right of the conditioning bar in a conditional probability expression.

3.3 conditioning profile

The DNA profile of an individual assumed to be a contributor to the DNA (or DNA profile) obtained from an evidentiary item in both propositions, H_1 and H_2 , of a pair of propositions.

3.4 evaluative analysis

An analysis using propositions that are used with the intent of producing an *LR* for reporting.

3.5 hierarchy of propositions

An organizational structure for propositions. Proposition pairs are classified by the level of information they can provide to the trier of fact: offense (e.g., “Mr. X raped V”), activity (e.g., “Mr. X had intercourse with V”), source (e.g., “The semen came from Mr. X”), sub-source (e.g., “Mr. X is a contributor to this DNA.”), and sub-sub-source (e.g., “Mr. X is the minor contributor to this DNA mixture”). Further explanations can be found in Buckleton et al. [2], pp. 46-48.

3.5.1 sub-source level proposition

A proposition that specifies a defined set of assumed contributors (known or not) to the questioned DNA (evidentiary) sample.

3.5.2

sub-sub-source level proposition

A proposition that specifies an assumed contributor (known or not) to one contribution of a questioned DNA (evidentiary) mixture (e.g., the major contributor of a two-person mixture).

3.6

intimate contributor

An individual from whose body a biological evidentiary item has been directly obtained.

3.7

probabilistic genotyping

The use of biological modeling (i.e., statistical modeling informed by biological data), statistical theory, computer algorithms, and/or probability distributions, to infer genotypes and/or calculate likelihood ratios.

3.8

probabilistic genotyping system

Software, or software and hardware, which utilizes a probabilistic genotyping approach to infer genotypes and/or calculate likelihood ratios.

3.9

proposition

A statement that is true or false, associated with the standpoint of one of the parties on a disputed issue of interest.

4 Requirements

Refer to Annex A, Best Practice Recommendations, for additional normative information on the following requirements.

4.1 The laboratory shall have a documented policy that outlines the difference between relevant and irrelevant information.

4.2 The laboratory shall have a documented policy on the reporting of results within the hierarchy of propositions. The policy shall define the level within the hierarchy of propositions (e.g., sub-source or sub-sub-source) that the laboratory reports.

4.3 The laboratory shall have a documented policy defining when a conditioning profile will be used. Support for the assumption of non-intimate conditioning contributors shall be documented in the case file.

4.4 Where multiple persons of interest (POIs) have LRs that support an association to a DNA mixture, within the capabilities of the approach used, an analysis shall be performed using proposition pairs that test whether the multiple POIs can be included together in the observed DNA profile.

4.5 If an analysis requires the number of contributors to be declared for each proposition or pair of propositions tested, then the laboratory shall have a documented policy regarding the assignment of the number of contributors for each proposition or pair of propositions.

4.5.1 If an analysis requires a predefined number of contributors, then an assessment of the number of contributors shall be documented in the case file prior to comparison to any individuals who are not intimate or reasonably assumed (e.g., consensual partners) contributors.

4.5.2 If the approach used to calculate the LR cannot accommodate different numbers of contributors in H_1 and H_2 , then the profile of the POI shall not be used in the initial assignment of the number of contributors.

4.5.3 Use of conditioning profiles in the contributor number assessment shall be documented in the case file.

4.5.4 If the laboratory allows for drop-in events in their probabilistic genotyping procedures, the laboratory shall have a documented policy regarding contributor number assumptions that require drop-in events.

4.5.5 Any reassessment of the number of contributors based upon a probabilistic genotyping system's diagnostic calculations, or subsequent interpretations or LR calculations, shall be documented in the case file. This documentation shall include the reason for the change.

4.6 The laboratory shall have a documented policy on when and how propositions will include a relative of a conditioned donor within the set of unknowns.

4.7 Reporting only one, or a subset, of the LRs for evaluative analyses that have been conducted for multiple proposition pairs, differing by the conditioning profiles and/or the number of contributors, shall require documentation in the case file providing the reason the selected LRs were chosen.

4.7.1 Lacking any direction from either party, the laboratory shall formulate one or more reasonable proposition pairs that address each party's assumed interests.

4.7.2 The laboratory shall reassess the data under a new pair of propositions if presented with a reasonable request (i.e., supported by the associated data, and presented in a timely manner) from either party regarding an alternate proposition pair.

Annex A (normative)

Best Practice Recommendations

A.1 Within the capabilities of the analysis approach used, the laboratory should report results for a pair of propositions that addresses the issue of interest. This level is the highest level in the hierarchy for which the forensic scientist can provide information. Hence, the laboratory should report results at sub-source level rather than sub-sub-source level.

A.2 If a single source profile is deduced from the mixture, even where sub-source propositions for mixtures are possible, then a sub-sub-source proposition pair may be used. However it is permissible, and may be advisable, to continue to operate at the sub-source level.

A.3 A profile should be assigned as a conditioning profile to a mixture when an individual is identified as an intimate contributor, or when it is reasonable to assume their presence based on case specific information, and the associated data supports the assumption. The conditioning profile could be from the complainant, POI, or other individual depending on the case scenario.

A.4 When conditioning on individuals in non-intimate samples, propositions that do not condition upon that individual's profile may also be examined. A positive $\log(LR)$ value for an individual may be used as evidence of support for calculations that condition upon them.

A.5 The analysis should separate the propositions into their simplified constituents (i.e., simple proposition pairs²) when an LR favoring H_1 has resulted from a compound proposition pair³ incorporating multiple POIs under H_1 and none of the POIs under H_2 , in order to establish the weighting and the consequent probative value of the evidence per contributor under H_1 .

A.5.1 The results of the analyses using the simple proposition pairs should be included in the report unless the compound conclusion supports exclusion.

A.5.2 A simple proposition pair may include conditioning upon each POI separately in H_1 , where all other contributors are treated as unknowns. This could, however, contradict the assertions of the party represented by that proposition.

A.5.3 A simple proposition pair may include conditioning upon multiple POIs in H_1 and one or a subset of POIs in H_2 with the goal of isolating the weight attributable to the POI not conditioned upon in H_2 . In the case of non-intimate samples, if a POI's profile is assumed in both H_1 and H_2 , then an LR shall also be calculated for a simple proposition pair with the POI not assumed in H_2 .

A.6 While the number of contributors to an evidentiary sample is strictly unknown, assessments of the number of contributors should involve profile traits such as

- a) The number of clearly defined alleles;
- b) The relative intensities of the results (e.g., relative peak heights);

² **simple proposition pair**

A pair of propositions where no more than one POI in H_1 is replaced with an unknown donor in H_2 or vice versa.

³ **compound proposition pair**

A pair of propositions where more than one POI in H_1 is replaced with unknown donors in H_2 or vice versa.

- c) The relative proportions of peaks in expected artifact positions (e.g., STR reverse stutters);
- d) Any assumptions regarding the contribution of a conditioning profile's alleles; and
- e) Any assumptions regarding rare genetic traits (e.g., tri-allelic genotypes).

Probabilistic methods may be used to provide estimated probabilities for the number of contributors.

A.6.1 A number of alternatives for the number of contributors may be used.

A.6.2 The number of contributors may be the same or different for H_1 and H_2 .

A.6.3 For approaches that utilize probabilities, as opposed to probability densities, the number of contributors may be assigned values that maximize the probability of the evidence separately under H_1 and H_2 . In such cases, the number of contributors under each proposition need not be the same within a given likelihood ratio.

A.6.4 Assigning the number of contributors may involve assessing possible drop-in events and also assessing or specifying the number of drop-in events allowed. Typically, if more than two drop-in events would be required, the assumed number of contributors should be increased [9, 11].

A.7 Investigative analyses⁴ should not be reported but retained in the case file. The report may include a comment that investigative analyses have been undertaken.

⁴ **investigative analysis**

An analysis using propositions that are used to determine the best pair(s) to use for an evaluative analysis.

Annex B (informative)

Bibliography

- 1] Buckleton, John, Bright, Jo-Anne, Taylor, Duncan, Evett, Ian, Hicks, Tacha, Jackson, Graham, and Curran, James. *Helping formulate propositions in forensic DNA analysis*. Science and Justice, 2014.
- 2] Buckleton, John, Bright, Jo-Anne, and Taylor, Duncan. *Forensic DNA Evidence Interpretation, 2nd ed.* CRC Press, Boca Raton, 2016.
- 3] Buckleton, John and Curran, James. *A discussion of the merit of random man not excluded and likelihood ratios*. Forensic Science International: Genetics, 2008.
- 4] Buckleton, John, Curran, James, and Gill, Peter. *Towards understanding the effect of uncertainty in the number of contributors to DNA stains*. Forensic Science International: Genetics, 2007.
- 5] Cook, R., Evett, Ian, Jackson, Graham, Jones, P., and Lambert, J. *A hierarchy of propositions: deciding which level to address in casework*. Sci Justice 1998.
- 6] European Network of Forensic Science Institutes, *ENFSI guideline for evaluative reporting in forensic science, version 3.0*, 2015.
- 7] Evett, Ian, Jackson, Graham, and Lambert, J. *More on the hierarchy of propositions: exploring the distinction between explanations and propositions*. Science & Justice 2000.
- 8] Gill, Peter and Haned, Hinda. *A new methodological framework to interpret complex DNA profiles using likelihood ratios*. Forensic Science International: Genetics, 2013.
- 9] Gill, Peter, Kirkham, A., and Curran, James, *LoComatioN: A software tool for the analysis of low copy number DNA profiles*. Forensic Science International, 2007.
- 10] Gittelsohn, Simone, Kalafut Tim, Myers, Steven, Taylor, Duncan, Hicks, Tacha, Taroni, Franco, Evett, Ian, Bright, Jo-Anne, and Buckleton, John. *A practical guide for the formulation of propositions in the Bayesian approach to DNA evidence interpretation in an adversarial environment*. Journal of Forensic Science, 2016.
- 11] Haned, Hinda, Slooten, Klaas, and Gill, Peter. *Exploratory data analysis for the interpretation of low template DNA mixtures*. Forensic Science International: Genetics, 2012.
- 12] Hicks, Tacha, Biedermann, Alex, de Koeijer, Joziás., Taroni, F., Champod, Christophe, and Evett, Ian. *The importance of distinguishing information from evidence/observations when formulating propositions*. Science and Justice 2015.
- 13] Federal Bureau of Investigation, *Quality Assurance Standards for Forensic DNA Testing Laboratories*, available at www.fbi.gov and www.swgdam.org
- 14] References at www.swgdam.org