Forensic Science International: Genetics 25 (2016) 191–197

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

DNA Commission of the International Society for Forensic Genetics: Recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications

M.D. Coble[a,*], J. Buckleton[b,c], J.M. Butler[d], T. Egeland[e], R. Fimmers[f], P. Gill[g,h], L. Gusmão[i,j,k], B. Guttman[l], M. Krawczak[m], N. Morling[n], W. Parson[o,p], N. Pinto[j,k,q,r], P.M. Schneider[s], S.T. Sherry[t], S. Willuweit[u], M. Prinz[v]

[a] National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, MD, USA
[b] ESR, Private Bag 92021, Auckland 1142, New Zealand
[c] National Institute of Standards and Technology, Statistical Engineering Division (Guest Researcher), Gaithersburg, MD, USA
[d] National Institute of Standards and Technology, Special Programs Office, Gaithersburg, MD, USA
[e] Norwegian University of Life Sciences, Oslo, Norway
[f] Institute for Medical Statistics, Informatics, and Epidemiology, University Bonn, Germany
[g] Norwegian Institute of Public Health, Oslo, Norway
[h] University of Oslo, Oslo, Norway
[i] State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil
[j] IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, Portugal
[k] Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal
[l] National Institute of Standards and Technology, Software and Systems Division, Gaithersburg, MD, USA
[m] Institute of Medical Informatics and Statistics, Christian-Albrechts University of Kiel, Germany
[n] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark
[o] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[p] Forensic Science Program, The Pennsylvania State University, PA, USA
[q] Institute for Research and Innovation in Health (I3S), University of Porto, Porto, Portugal
[r] Centre of Mathematics of the University of Porto, Porto, Portugal
[s] Institute of Legal Medicine, Faculty of Medicine, University of Cologne, Germany
[t] National Center for Biotechnology Information, National Library of Medicine, NIH, Bethesda, MD, USA
[u] Department of Forensic Genetics, Institute of Legal Medicine and Forensic Sciences, Charité—Universitätsmedizin, Berlin, Germany
[v] John Jay College of Criminal Justice, New York, USA

ARTICLE INFO

Article history:
Received 29 August 2016
Accepted 2 September 2016
Available online 4 September 2016

Keywords:
Biostatistical software
Software validation
Validation and verification
Forensic genetics
Software training and testing

ABSTRACT

The use of biostatistical software programs to assist in data interpretation and calculate likelihood ratios is essential to forensic geneticists and part of the daily case work flow for both kinship and DNA identification laboratories. Previous recommendations issued by the DNA Commission of the International Society for Forensic Genetics (ISFG) covered the application of bio-statistical evaluations for STR typing results in identification and kinship cases, and this is now being expanded to provide best practices regarding validation and verification of the software required for these calculations. With larger multiplexes, more complex mixtures, and increasing requests for extended family testing, laboratories are relying more than ever on specific software solutions and sufficient validation, training and extensive documentation are of upmost importance.

Here, we present recommendations for the minimum requirements to validate bio-statistical software to be used in forensic genetics. We distinguish between developmental validation and the responsibilities of the software developer or provider, and the internal validation studies to be performed by the end user. Recommendations for the software provider address, for example, the documentation of the underlying models used by the software, validation data expectations, version control, implementation and training support, as well as continuity and user notifications. For the internal validations the recommendations include: creating a validation plan, requirements for the range

* Corresponding author at: National Institute of Standards and Technology, Applied Genetics Group, 100 Bureau Drive MS 8314, Gaithersburg, MD 20899-8314, USA.
E-mail address: mcoble@nist.gov (M.D. Coble).

http://dx.doi.org/10.1016/j.fsigen.2016.09.002
1872-4973/Published by Elsevier Ireland Ltd.

of samples to be tested, Standard Operating Procedure development, and internal laboratory training and education. To ensure that all laboratories have access to a wide range of samples for validation and training purposes the ISFG DNA commission encourages collaborative studies and public repositories of STR typing results.

Published by Elsevier Ireland Ltd.

1. Introduction

Forensic genetics is experiencing an increase in data volume and complexity, and the interpretation of these data is becoming more and more dependent upon the use of appropriate bio-statistical computer programs. Software for calculating likelihood ratios to evaluate trace evidence or competing kinship scenarios has been in use for many years now, and several groups have described validation exercises of either in-house, open source, or commercial software packages [1–15].

These publications vary notably in terms of the validation approach taken, and standardized reporting of which quality measures were invoked, which tests have been successfully completed, and which software documentation was available. This information is not only of interest to the forensic scientist but also to the legal community. For quality measures, a distinction must be drawn between the responsibility of the software developer or provider, e.g. for code review, version control, documentation of the underlying theory and validation against known data sets, and the responsibility of the end user, e.g. internal validation under local laboratory conditions, formulation of standard operating procedures (SOPs), and training and competency testing.

International industry standards apply to software validation, verification [16] and test documentation [17]. These standards can be simplified and extrapolated [18] to forensic genetics. For internal validation, the goal is similar to other analysis methods: to test the proper function and assess accuracy and limitations of the methods. Previous recommendations on forensic method validation and application of genetic analyses are useful to be read in conjunction with these guidelines [19–25].

The International Society for Forensic Genetics (ISFG) has convened a DNA Commission to establish validation guidelines for bio-statistical software to be used in forensic genetics. Examples include software to calculate statistics for: single-source samples, autosomal DNA mixtures of two or more individuals with no drop-out, or where drop-out and drop-in are possible, paternity and kinship testing, and haploid marker interpretation. The goal of the DNA Commission was to carve out a consensus view on the minimum requirements for the validation (is it doing the right thing?) and verification (is it doing the thing right?) of a software program (V&V) [16] and to describe the software test documentation (STD) [17] to be generated by the software provider. The DNA Commission differentiated developmental from internal (laboratory) validation and emphasizes that the software used is an integral part of the evidential process and should not be treated as a separate and isolated component.

2. Provider responsibilities and developmental validation

The software developer has the burden to specify and document the assumptions and genetic/statistical models underlying the software program and refer to mathematical/statistical proofs or provide these with the software. Prior to promoting their software for practical use, the provider or developer must conduct a developmental validation demonstrating that the intended calculations are being performed correctly and that they provide the expected results. The data sets used for validation should be made publicly available alongside the validation results, as is outlined below.

2.1. Underlying models and developer's validation

Recommendation 1

Bio-statistical software for forensic genetic applications should be accompanied by scientific papers or information or guidance materials, such as a user manual, describing the underlying method. The population genetic and data model(s) used should be explicitly described and disclosed to allow the reproducibility of all the computations by other means (algebraic formulae, other software programs or statistical approaches) as publication in peer-reviewed journals

The DNA Commission encourages software providers or developers to report the theoretical assumptions underlying their product or refer to already published models. We also encourage the publication of the design and outcome of their developmental validation in peer-reviewed journals. We discourage insufficiently documented or described software where the end user cannot adequately explain to the trier of fact (e.g. judge or jury) the theoretical basis of the software used.

Recommendation 2

Bio-statistical software for forensic genetic applications should be validated according to particular requirements and specific intended use. The software developer's validation should use publicly-available data sets or disclose the used data set otherwise. The result of the software developer's validation and its environment (hardware and software dependencies) should be documented and disclosed

One of the principles of scientific research is that any new finding should be amenable to independent replication. The DNA Commission therefore encourages software providers or developers to verify and validate their software (e.g. by generating or using validation data sets with known outcomes) along with the parameters necessary for the software to work (e.g. population allele counts for frequency calculation). Verification may be assessed using code review. This information could then be publicized so as to support interested laboratories with their own internal training and explorative testing of the software.

The test cases of the validation data should be designed so as to cover all of the software functionality, to be complex enough to detect installation errors, and to be generic enough to also serve as a basis for testing the consistency of future versions of the validated software. Although the goal of internal validation is not to repeat developmental validation, making the data and parameters used for the latter publicly available may add extra benefit in that it would allow laboratories to investigate the local performance of the software under the conditions of the developmental validation, if they so wish. Validation test results should be documented (and disclosed) following a test plan [17] as well as system requirements and platform (hardware and software) specification.

The validity of the results obtained from a given validation data set should also be assessed by way of comparison to the results obtained through hand calculations of algebraic formulae (if possible), using alternative statistical approaches where applicable