Transfusion and Apheresis Science 52 (2015) 35–41

Contents lists available at ScienceDirect

Transfusion and Apheresis Science

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

Review

Evaluation of the measurement uncertainty in screening immunoassays in blood establishments: Computation of diagnostic accuracy models

Paulo Pereira [a,*], James O. Westgard [b], Pedro Encarnação [c], Jerard Seghatchian [d,**]

[a] Department of Quality Assurance, Portuguese Institute of Blood and Transplantation, Avenida Miguel Bombarda 6, Lisbon 1000-208, Portugal
[b] Department of Pathology and Laboratory Medicine, University of Wisconsin Medical School, Madison, WI, United States
[c] Católica Lisbon School of Business and Economics, Catholic University of Portugal, Lisbon, Portugal
[d] International Consultancy in Blood Components Quality/Safety Improvement, Audit / Inspection and DDR Strategy, London, United Kingdom

ARTICLE INFO

Keywords:
Blood establishment
Diagnostic accuracy
Measurement uncertainty
Risk assessment
ROC

ABSTRACT

The European Union regulation for blood establishments does not require the evaluation of measurement uncertainty in virology screening tests, which is required by ISO 15189 guideline following GUM principles. GUM modular approaches have been discussed by medical laboratory researchers but no consensus has been achieved regarding practical application. Meanwhile, the application of empirical approaches fulfilling GUM principles has gained support. Blood establishments' screening tests accredited by ISO 15189 need to select an appropriate model even GUM models are intended uniquely for quantitative examination procedures. Alternative (to GUM) models focused on probability have been proposed in medical laboratories' diagnostic tests. This article reviews, discusses and proposes models for diagnostic accuracy in blood establishments' screening tests. The output of these models is an alternative to VIM's measurement uncertainty concept. Example applications are provided for an anti-HCV test where calculations were performed using a commercial spreadsheet. The results show that these models satisfy ISO 15189 principles and that the estimation of clinical sensitivity, clinical specificity, binary results agreement and area under the ROC curve are alternatives to the measurement uncertainty concept.

Contents

* Corresponding author. Department of Quality Assurance, Portuguese Institute of Blood and Transplant, Avenida Miguel Bombarda 6, 1000-208 Lisboa, Portugal. Tel.: +351 210063047; fax: +351 210063070.
E-mail address: paulo.pereira@ipst.min-saude.pt (P. Pereira).
** Corresponding author.
E-mail address: jseghatchian@btopenworld.com (J. Seghatchian).

http://dx.doi.org/10.1016/j.transci.2014.12.017
1473-0502/

1. Introduction

Tests results, with an ordinal quantity (entry 1.26 of ref. 1), in medical laboratories are commonly referred to as "diagnostic tests" [2] or "qualitative tests" [3]. In blood establishments they are known as "screening tests" [4]. In screening immunoassays, the measurement on an ordinal scale is compared to a clinical decision level or cutoff to produce a binary test result (positive/negative).

tj;8Since the International Organization for Standardization (ISO)/the International Electrotechnical Commission (IEC) 17025 publication in 1999, the expression of measurement uncertainty is a technical requirement, within this accreditation program in global general laboratory. The International Vocabulary of Metrology (VIM) defines "measurement uncertainty" as the "non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used" (2.26 of ref. 5). The definition limits its application to numerical quantity results; however medical laboratories already dealt with ordinal binary results (note: these results are also entitled "qualitative" or "semi-quantitative" results). ISO 15189 first edition stated that medical laboratories "shall determine the uncertainty of results, where relevant and possible" [6]. The current ISO 15189 edition (2012) requires its determination and also recommends that laboratories define the performance requirements for measurement uncertainty and regularly review their estimates of measurement uncertainty [7]. In contrast, the European Union directives [8–11] and the American Association of Blood Banks standards [12] do not require the determination of measurement uncertainty for blood establishments' screening tests results (note: in the United States the blood establishments' laboratories must fulfil Clinical Laboratory Improvement Amendments requirements, which do not require determination of measurement uncertainty).

Despite the global metrology organizations, such as the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), as well as the ISO, recommend that measurement uncertainty shall be determined with GUM methods, ISO 15189 does not request its determination fulfilling GUM principles [5]. Also in the medical laboratory field, The Australian National Pathology Accreditation Advisory Council (NPAAC) guideline (2007) recommends a set of GUM empirical approaches [13], and the Clinical and Laboratory Standards Institute (CLSI) C51-A guideline (2012) features GUM modular and empirical approaches [14]. However, all the models are intended only for quantitative tests. The measurement uncertainty could be applied as part of the total quality management system mainly on analyse and control stages [15].

CLSI defines "diagnostic accuracy" as the "the extent of agreement between the information from the test under evaluation and the diagnostic accuracy criteria" (entry 5.3 of ref. 3) or "the ability of a test system to obtain the correct result" (entry 4.1 [16]). It can be estimated by sensitivity and specificity pairs, likelihood ratio of positive and negative result pairs, and area under the receiver operating characteristic (ROC) curve. This paper discusses diagnostic accuracy models, relating the uncertainty to the intended use of examination results, i.e., post-transfusion safety.

Despite Fuentes-Arderiu [17] and Dybkaer [18] having proposed terminologies suitable for ordinal quantity tests, they are not used in common practice, for what this paper adopts the terminology particular to the presented models.

The theoretical principles were implemented using standard spreadsheet software (Microsoft® Excel® 2013). The spreadsheet validation was done according to the U.S. Food and Drug Administration (FDA) recommendations [19]. The spreadsheets supporting the results of this article are included within the article as supplementary material. They were intended to perform modelling for this paper to allow easy handling and automation of calculation. They are not suited for use in blood establishments.

2. Methods and materials

2.1. Diagnostic accuracy models when the diagnosis is known

Diagnostic accuracy methods measure the agreement between the screening test binary results and the diagnostic accuracy criteria (i.e., disease/non-disease). A Bayesian probability framework is adopted. The patients sample should be carefully selected to prevent "spectrum bias", i.e., "bias between estimated test performance and true test performance when the sample used for evaluating an assay does not properly represent the entire disease spectrum over the target (intended-use) population" (entry 4.2 of ref. 2). When the diagnosis is unknown, a comparative method should be used to measure the degree of concordance between the screening test binary results and the binary results of a comparative test [3].

The measurement of the percentage of true-positive results among the test results for a sample known to be positive for the test is known as "clinical sensitivity" (entry 5.3 of ref. 3) $se[\%]$ is measured through the mathematical model $[TP/(TP + FN)]\cdot 100$, where TP is the number of true-positive results and FN is the number of false-negative results. On the non-disease sample, the percentage of true-negative results among the test results for a sample known to be negative for the test is known as "clinical specificity" (entry 5.3 of ref. 3) $sp[\%]$ which is measured through the model $[TN/(FP + TN)]\cdot 100$, where TN is the number of true-negative results and FP is the number of false-positive results. Transfusion safety requires that blood establishments' immunoassays should have a high sensitivity to ensure that results from infected donors have a high probability to be