TECHNICAL SPECIFICATION

ISO/TS 27878

First edition
2023-01

Reproducibility of the level of detection (LOD) of binary methods in collaborative and in-house validation studies

Reproductibilité de la limite de détection (LD) des méthodes binaires pour des études de validation internes et collaboratives

Reference number
ISO/TS 27878:2023(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

An appropriate approach for the validation of binary methods will often differ considerably from that of quantitative methods. Nevertheless, core concepts from the validation of quantitative methods can be successfully carried over to binary methods. In particular, the precision of a method – a performance characteristic usually associated with quantitative methods – can be determined for the level of detection (LOD) of binary methods.

In analytical chemistry, one of the fundamental indicators of method performance is the reproducibility of quantitative test results as described in ISO 5725 (all parts)[1]. This aspect of method performance is not usually taken into consideration in the validation of binary methods. However, in the last few years, novel validation approaches have been proposed in which the reproducibility of a binary method can be determined and meaningfully interpreted.

Why is it important to determine a method's reproducibility? In order to answer this question, consider an example from the field of microbiology. Take the case that, in the validation study, a method's LOD is determined as 3 CFU/ml (CFU = colony forming unit), but that the LOD is sometimes much higher depending on the laboratory or on the test conditions. Failing to detect the occasional unreliability of the method could lead to mistakes in routine laboratory determinations. On the other hand, if an LOD of 300 CFU/ml is obtained in the validation study, the method will not be validated even though this excessive LOD is not representative of its average performance. Accordingly, both the average LOD value and the reproducibility parameter – describing the variability of the LOD across laboratories or test conditions – capture important information about the performance of the method and should be determined in the course of the validation process.

In order to accomplish this, a suitable approach should be identified for the conversion of the binary results into quantitative ones. In this standard, two parametric models for the calculation of the LOD will be used: one model for methods for discrete measurands, e.g. microbiological and Polymerase Chain Reaction (PCR) methods, and one model for methods for continuous measurands, e.g. chemical methods.

Two different study designs will be applied. In the conventional approach, test conditions vary randomly from one laboratory to the other, whereas in the factorial approach, at least to some extent, test conditions are controlled. The factorial approach makes it possible to assess different sources of errors such as the variability arising in connection with different analysts, different instruments, different lots of reagents, different elapsed assay times, different assay temperatures etc. Such an approach also allows a reduction in workload and fewer participating laboratories.

Reproducibility of the level of detection (LOD) of binary methods in collaborative and in-house validation studies

1 Scope

This document provides statistical techniques for the determination of the reproducibility of the level of detection for

a) binary (qualitative) test methods for continuous measurands, e.g. the content of a chemical substance, and

b) binary (qualitative) test methods for discrete measurands, e.g. the number of RNA copies in a sample.

The reproducibility precision is determined according to ISO 5725 (all parts).

Precision estimates are subject to random variability. Accordingly, it is important to determine the uncertainty associated with each estimate, and to understand the relationship between this uncertainty, the number of participants and the experimental design. This document thus provides not only a description of statistical tools for the calculation of the LOD reproducibility precision, but also for the standard error of the estimates.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ISO 5725-1, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534-1 and ISO 5725-1 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
factor
binary or quantitative parameter within the method that can be varied at two or more levels within the limits of the specified method

EXAMPLE Technician.

3.2
factor level
value of the *factors* (3.1) within the experimental design

EXAMPLE Technician 1, Technician 2, etc.

3.3
level of detection
LOD
concentration from which on the *POD* (3.4) is not below a specified limit, e.g. 0,5 or 0,95 ($LOD_{50\%}$ or $LOD_{95\%}$).

Note 1 to entry: This definition is mathematically equivalent to the definitions for "level of detection" in ISO 16140-1[2], ISO 16140-2[3] and ISO 16140-4[4]. It differs from the definition used for chemical and physical methods for which a "limit of detection" is defined as the lowest quantity of an analyte that can be distinguished from the absence of that analyte with a stated confidence level.

Note 2 to entry: In this document, the term concentration (or concentration level) is used as a generic term to mean not only the actual concentration in the case of a measurand that can be quantified on a continuous scale, but also the number of colony forming units or DNA copies per aliquot in the case of measurands which are quantified on a discrete scale.

3.4
probability of detection
POD
probability of a positive analytical outcome of a qualitative test method at a given concentration for a specific sample type

Note 1 to entry: This definition is based on the two slightly different definitions for "probability of detection" in ISO/TS 16393[6] and ISO 16140-1, ISO 16140-2 and ISO 16140-4.

Note 2 to entry: The POD is a measure of the probability of a positive analytical result and thus a theoretical value which can be approximated by a mathematical model.

3.5
rate of detection
ROD
proportion of positive analytical outcomes in a test series, when a qualitative method is performed several times with a specific sample

Note 1 to entry: The ROD is not a theoretical value based on a mathematical model [like the *POD* (3.4)] but the result of a series of repeated tests performed on a given sample.

4 Symbols

p	number of participating laboratories
σ_L^2	between-laboratory variance
POD = P	probability of detection
x	concentration level (see Note 1 to entry 3.3) at which the POD is calculated
ROD	rate of detection
$LOD_{50\%} = L_{50}$	50 % of the level of detection
$LOD_{95\%} = L_{95}$	95 % of the level of detection
L, H, B, C	global model parameters for the four-parameter sigmoid curve

a_i laboratory-specific correction of laboratory i for the global inflection point C

$N\left(\mu,\sigma^2\right)$ normal distribution with mean μ and variance σ^2

5 General principles

5.1 General considerations

In order to ensure that tests are conducted in the same manner in all participating laboratories, the test method should be standardized. All tests forming part of an experiment within an individual laboratory or of an interlaboratory experiment shall be carried out according to the corresponding standardized protocol.

The statistical methods described in this document are applicable for binary test methods which yield a yes/no result (e.g. the substance of interest is present or absent). For such test methods, one of the main criteria of the method's fitness for purpose is the level of detection (e.g. $LOD_{50\%}$ or $LOD_{95\%}$), i.e. the (concentration) level required to ensure a POD of 50 % or 95 %. The aim is thus to determine LOD values for the individual laboratories as well as an overall LOD across laboratories. The precision of the method can then be evaluated in terms of the variability to which the laboratory-specific LOD values are subjected.

The laboratory-specific LOD values and the mean LOD across laboratories can be computed based on a mathematical model for the relationship between level, x, and probability of detection $\text{POD}_i\left(x\right)=P_i\left(x\right)$ for laboratory i: The $LOD_{95\%}$ of laboratory i is then the lowest level, x, for which $\text{POD}_i\left(x\right)=P_i\left(x\right)\geq 0{,}95$.

5.2 Considerations for the conventional approach

The conventional approach is based on the assumption that, according to the design used in ISO 5725-2, all tests are performed under repeatability conditions in each of the laboratories involved. In particular, all tests in the laboratory are performed by the same technician, with the same equipment, under the same conditions and directly one after the other. Test results are considered to have been obtained from different laboratories under reproducibility conditions, i.e. many factors contribute to observed variability, e.g. differences in equipment, environmental conditions, reagent batches or technician.

NOTE Validation protocols according to the conventional approach based on LOD and POD can be found in ISO 16140-2, ISO 16140-4 and ISO/TS 16393 and AOAC Guidelines[7]. Examples and further protocols are discussed e.g. in References [8][9][10][11][12] and [13].

5.3 Considerations for the factorial approach

Compared to the conventional approach, in which tests are made under repeatability conditions in each of the laboratories, the factorial approach systematically varies one or more factors. For instance, half the tests are conducted with reagents from batch A, and the other half with reagents from batch B. Thus, the factorial approach makes it possible to ensure the full spectrum of test conditions is covered in the validation study and assess contributions to variability from separate sources of error. This approach translates to more efficient and reliable estimation of the total variability.

NOTE Validation protocols based on LOD for microbiological methods according to the factorial approach are given in ISO 16140-4[4] and ISO 16140-5[5].

6 Conventional approach

6.1 Experimental design

Results from at least 8 participants, 4 concentration levels, and 8 replicates per level and laboratory are required to obtain a statistically reliable POD curve. However, with such a design, the reliability of the results may not be sufficient and will need to be checked. For more reliable estimation of the LOD and the corresponding variability, it is recommended that results from at least 8 participants, 5 concentration levels, and 12 replicates per level and laboratory are available. If the number of participants is increased, the number of replicates can be reduced.

The lowest concentration level should be selected so that no further reduction in POD is expected, even if the concentration level is further reduced. The highest concentration level should be selected in such a way that no further increase in POD is to be expected even if the concentration level is further increased. The expected proportions of positive test results across laboratories should be between 20 % and 80 % for at least two concentration levels.

The proportion of positive test results expected at the beginning of the collaborative trial usually differs from the final POD. This may mean that the proportion of positive test results actually determined in the collaborative trial does not meet the above requirements. In this case, the results of the evaluation and, in particular, the calculated reproducibility of the LOD can only be regarded as an estimate.

NOTE These recommendations for the experimental design are based on simulation studies in which the standard error of the estimate of the laboratory standard deviation was evaluated.

6.2 Statistical model for methods for continuous measurands

The calculation of the LOD is based on a generalized linear mixed-effects model (GLMM) together with a four-parameter sigmoid curve given by Formula (1):

$$\mathrm{POD}_i = P_i = \frac{L-H}{1+\left(\frac{x}{a_i C}\right)^B} + H \tag{1}$$

where

i denotes the laboratory (i = 1, 2,..., p);

$\mathrm{POD}_i = P_i$ denote the probability of detection for laboratory i;

x denotes a given concentration level;

L, H, B, C are global model parameters (i.e. they are valid across all laboratories);

a_i denotes the laboratory-specific correction of laboratory i;

C denotes the global inflection point C.

It is assumed that the parameters, L (lowest probability of detection), H (highest probability of detection), and B (slope) are the same for all laboratories. The product $a_i C$ describes the location of the inflection point of the curve for laboratory i; for L = 0 %, H = 100 %, it corresponds to the concentration at which a POD of 50 % is reached. The value of this product is thus a direct measure of the performance of the specific laboratory. The parameter, C, corresponds to the performance of an average laboratory.

The a_i values are modelled as realizations of a random variable: It is assumed that the ln a_i values follow a normal distribution with

$$\ln a_i \sim N\left(0,\sigma_L^2\right)$$

The parameters *L, H, B, C* and σ_L^2 can be provided by maximum likelihood estimation, e.g. in mathematical-statistical software package. The variance σ_L^2 characterizes the variability of sensitivity between laboratories.

NOTE 1 Although there is no guarantee that the distribution of ln a_i values actually follows a normal distribution, the log transformation usually leads to a better approximation of the normal distribution. If the method displays poor precision, then the prediction range of the LOD values without log transformation could include infeasible negative values.

NOTE 2 It is assumed that the parameters *L, H, C* and *B* are the same for all laboratories, i.e. that the shape of the curve is sigmoidal and the same across laboratories. It should be checked whether this assumption is justified, e.g. through a graphic check of laboratory-specific POD curves.

The interpretation of the parameters will be explained with an example, see Reference [13]. A collaborative study of a method for the binary analysis of gluten in corn products was conducted to demonstrate that the binary test method can detect gluten contaminations below the threshold of 20 mg/kg gluten. A total of four corn sample lots with different gluten concentrations was submitted to 18 laboratories to evaluate the sensitivity and reproducibility of the test method. Each of the 18 laboratories conducted 10 tests for each of four concentration levels. Table 1 provides the corresponding numbers of positive results per laboratory and concentration level.

Table 1 — Number of positive test results per concentration level and laboratory (10 replicates)

Laboratory No.	Concentration level			
	0,88 mg/kg	**2,42 mg/kg**	**5,48 mg/kg**	**9,38 mg/kg**
01	0	10	10	10
02	0	10	10	10
03	0	10	10	10
04	0	10	10	10
05	0	10	10	10
06	0	10	10	10
07	0	10	10	10
08	0	9	10	10
09	0	10	10	10
10	0	9	8	10
11	0	10	10	10
12	0	10	10	10
13	0	10	10	10
14	0	10	10	10
15	0	9	10	10
16	0	10	10	10
17	0	10	10	10
18	2	10	10	10

Figure 1 shows the POD curve of a laboratory with average performance (solid line) along with 95 % prediction range of laboratory-specific POD (dark grey zone) and 95 % prediction range of laboratory-specific RODs (light grey step-functions). The numbers adjacent to the diamonds indicate the laboratory numbers having obtained the corresponding ROD.