INTERNATIONAL STANDARD

ISO 16140-4

First edition
2020-07

Microbiology of the food chain — Method validation —

Part 4: Protocol for method validation in a single laboratory

Microbiologie de la chaîne alimentaire — Validation des méthodes —

Partie 4: Protocole pour la validation de méthodes dans un seul laboratoire

Reference number
ISO 16140-4:2020(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 34, *Food products*, Subcommittee SC 9, *Microbiology*, in collaboration with the European Committee for Standardization (CEN) Technical Committee CEN/TC 463, *Microbiology of the food chain*, in accordance with the Agreement on technical cooperation between ISO and CEN (Vienna Agreement).

A list of all parts in the ISO 16140 series can be found on the ISO website.

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

0.1 The ISO 16140 series

The ISO 16140 series has been expanded in response to the need for various ways to validate or verify test methods. It is the successor to ISO 16140:2003. The ISO 16140 series consists of six parts with the general title, *Microbiology of the food chain — Method validation:*

— *Part 1: Vocabulary*;

— *Part 2: Protocol for the validation of alternative (proprietary) methods against a reference method*;

— *Part 3: Protocol for the verification of reference methods and validated alternative methods in a single laboratory*;

— *Part 4: Protocol for method validation in a single laboratory*;

— *Part 5: Protocol for factorial interlaboratory validation for non-proprietary methods*;

— *Part 6: Protocol for the validation of alternative (proprietary) methods for microbiological confirmation and typing procedures*.

ISO 17468 is a closely linked International Standard, which establishes technical rules for the development and validation of standardized methods.

In general, two stages are needed before a method can be used in a laboratory.

— The first stage is the validation of the method. Validation is conducted using a study in a single laboratory followed by an interlaboratory study (see ISO 16140-2, ISO 16140-5 and ISO 16140-6). In the case when a method is validated within one laboratory (as described in this document), no interlaboratory study is conducted.

— The second stage is method verification, where a laboratory demonstrates that it can satisfactorily perform a validated method. This is described in ISO 16140-3. Verification is only applicable to methods that have been validated using an interlaboratory study.

In general, two types of methods are distinguished: reference methods and alternative methods.

A reference method is defined in ISO 16140-1:2016, 2.59, as an "internationally recognized and widely accepted method". The note to entry clarifies that "these are ISO standards and standards jointly published by ISO and CEN or other regional/national standards of equivalent standing".

In the ISO 16140 series, reference methods include standardized reference (ISO and CEN) methods as defined in ISO 17468:2016, 3.5, as a "reference method described in a standard".

An alternative method (method submitted for validation) is defined in ISO 16140-1:2016, 2.4, as a "method of analysis that detects or quantifies, for a given category of products, the same analyte as is detected or quantified using the corresponding reference method". The note to entry clarifies that: "The method can be proprietary. The term 'alternative' is used to refer to the entire 'test procedure and reaction system'. This term includes all ingredients, whether material or otherwise, required for implementing the method.".

This document, ISO 16140-4, addresses validation within a single laboratory. The results are therefore only valid for the laboratory that conducted the study. In this case, verification (as described in ISO 16140-3) is not applicable. ISO 16140-5 describes protocols for non-proprietary methods where a more rapid validation is required or when the method to be validated is highly specialized and the number of participating laboratories required by ISO 16140-2 cannot be reached. This document and ISO 16140-5 can be used for validation against a reference method. This document (regarding qualitative and quantitative methods) and ISO 16140-5 (regarding quantitative methods only) can also be used for validation without a reference method.

The flow chart in Figure 1 gives an overview of the links between the different parts mentioned above. It also guides the user in selecting the right part of the ISO 16140 series, taking into account the purpose of the study and the remarks given above.

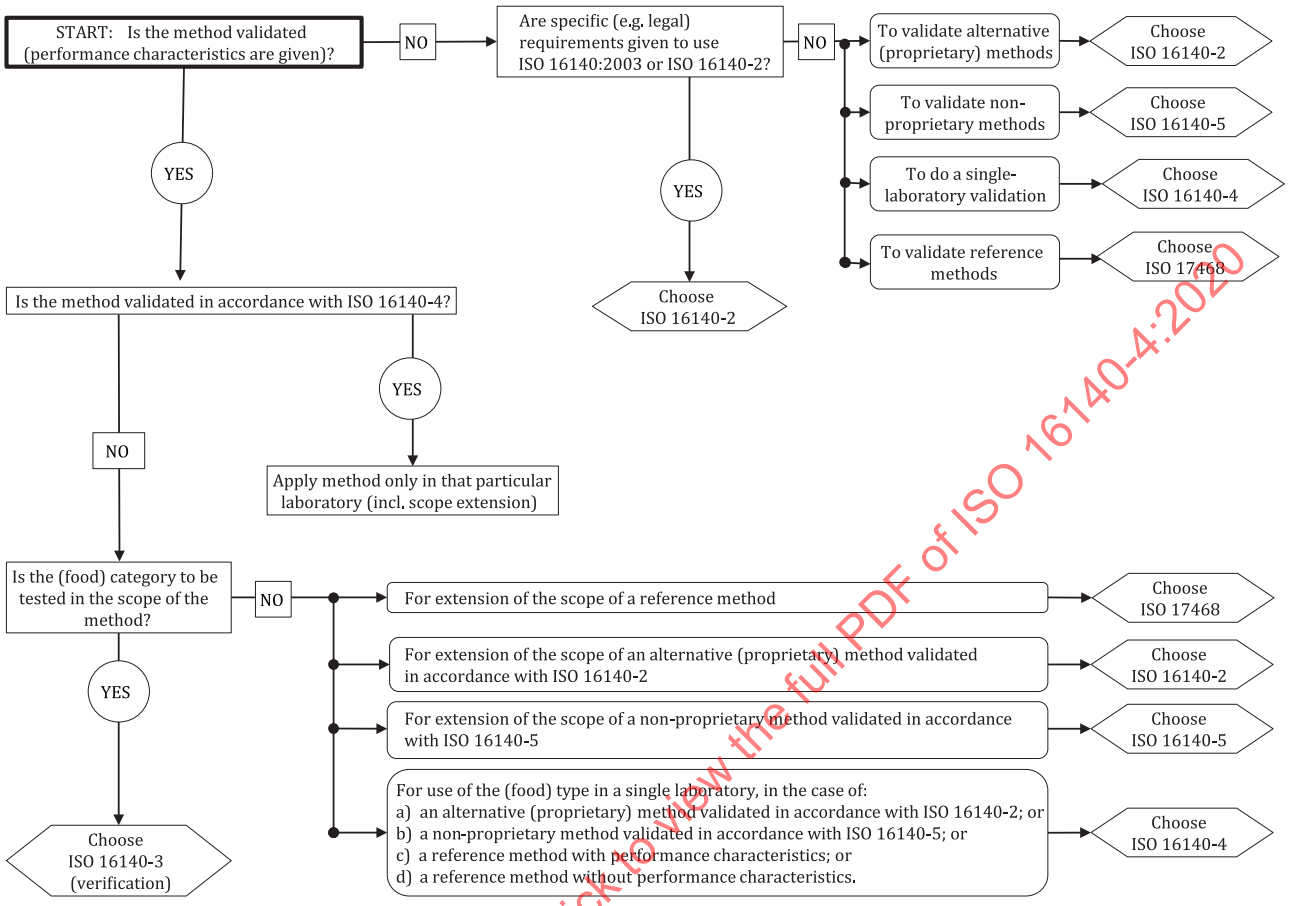


Figure 1 — Flow chart for application of the ISO 16140 series

NOTE In this document, the words "category", "type" and/or "item" are sometimes combined with "(food)" to improve readability. However, the word "(food)" is interchangeable with "(feed)" and other areas of the food chain as mentioned in Clause 1.

ISO 16140-6 is somewhat different from the other parts in the ISO 16140 series in that it relates to a very specific situation where only the confirmation procedure of a method is to be validated [e.g. the biochemical confirmation of *Enterobacteriaceae* (see ISO 21528-2)]. The confirmation procedure advances a suspected (presumptive) result to a confirmed positive result. The validation of alternative typing techniques (e.g. serotyping of *Salmonella*) is also covered by ISO 16140-6. The validation study in ISO 16140-6 clearly defines the selective agar(s) from which strains can be confirmed using the alternative confirmation method. If successfully validated, the alternative confirmation method can only be used if strains are recovered on an agar that was used and shown to be acceptable within the validation study. Figure 2 shows the possibilities where an alternative confirmation method validated in accordance with ISO 16140-6 can be applied (see text in the boxes).

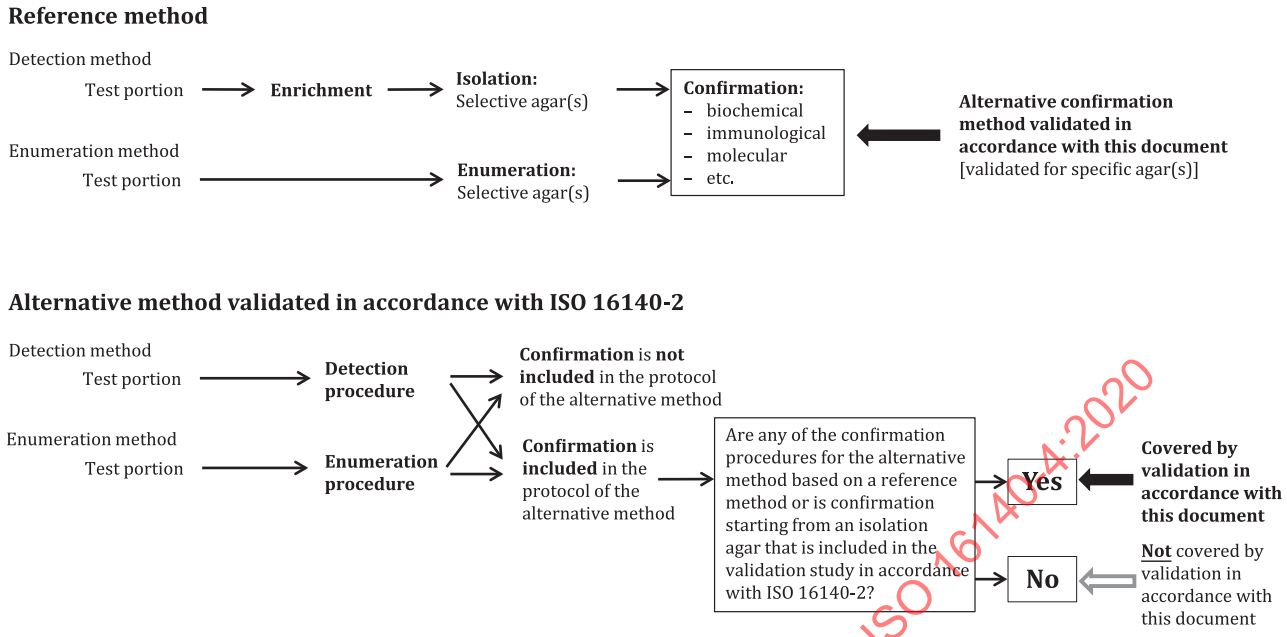


Figure 2 — Use of validated alternative confirmation methods (see ISO 16140-6)

EXAMPLE An example application of a validated alternative confirmation method is as follows.

An alternative confirmation method based on ELISA has been validated (in accordance with ISO 16140-6) to replace the biochemical confirmation for *Salmonella* as described in ISO 6579-1. In the validation study, XLD (mandatory agar in accordance with ISO 6579-1) plus BGA and a specified chromogenic agar (two optional agars for second plating in accordance with ISO 6579-1) were used as the agars to start the confirmation. The validated confirmation method can be used to replace the biochemical confirmation under the following conditions:

— by laboratories using the ISO 6579-1; or

— by laboratories using an ISO 16140-2 validated alternative method that refers to ISO 6579-1 for confirmation; or

— by laboratories using an ISO 16140-2 validated alternative method that starts the confirmation from XLD and/or BGA agar and/or the specified chromogenic agar.

The validated confirmation method cannot be used under the following conditions:

— by laboratories using an ISO 16140-2 validated alternative method that refers only to agars other than those included in the validation to start the confirmation (e.g. Hektoen agar and SS agar only); or

— by laboratories using an ISO 16140-2 validated alternative method that refers only to a confirmation procedure that does not require isolation on agar.

0.2 Validation protocols in the ISO 16140 series

An interlaboratory validation study, in accordance with ISO 16140-2, requires at least eight laboratories for quantitative methods and at least ten laboratories for qualitative methods. ISO 16140-5 is intended to be used for interlaboratory studies comprising four to seven laboratories for quantitative methods and four to nine laboratories for qualitative methods. ISO 16140-5 can only be used for non-proprietary methods. Table 1 provides an overview of the different protocols.

Table 1 — Overview of different validation protocols described in the ISO 16140 series

Number of participating laboratories	With reference method	Without reference method
1	This document: — factorial (see 5.1.1 and 5.2.1), or — conventional (see 6.1.1 and 6.2.1)	This document: — factorial (see 5.1.2 and 5.2.2), or — conventional (see 6.1.2 and 6.2.2)
4 to 7 (quantitative method)/ 4 to 9 (qualitative method)	ISO 16140-5: for non-proprietary methods only	ISO 16140-5: for non-proprietary quantitative methods only
≥ 8 (quantitative method)/ ≥ 10 (qualitative method)	ISO 16140-2: for the interlaboratory study part	Not applicable

The aim of this document is to assess the performance of detection or quantification methods within a single laboratory, typically across a number of (food) categories and (food) types. Single-laboratory validation of alternative methods for microbiological confirmation and typing procedures can also be performed under certain conditions: the general principles are the same as those described in ISO 16140-6 for the validation of alternative (proprietary) methods for microbiological confirmation and typing procedures (except there is no interlaboratory study). Further information is given in Annex G.

The protocols in this document only validate the method for the particular laboratory. A generalization to other laboratories is not within the scope of these protocols. However, extension to other laboratories is possible if this document is used as the first phase of validation of a reference method, to be followed by an interlaboratory study as described in ISO 17468.

If a reference method is available, the validation of a method is conducted by comparing the alternative method to the reference method. This allows inclusion of naturally contaminated samples in the validation process and thus provides a more realistic picture of the performance of the method. If no reference method is available, the validation process is based on samples with known contamination levels only. This document provides protocols for both situations.

The general principles for single-laboratory validations of detection and quantification methods are the same as those described in ISO 16140-2 for the validation of alternative (proprietary) methods against a reference method. This document cannot be used without ISO 16140-1 or ISO 16140-2, as many definitions and procedures are given in these International Standards. In addition to the validation parameters described in ISO 16140-2, this document describes the calculation of in-house repeatability and in-house reproducibility. Calculation of these parameters is not required if an interlaboratory study is to be conducted after the single-laboratory validation (i.e. if the single-laboratory validation is only the first phase of validation). Reliability of performance parameters obtained with this document is comparable to ISO 16140-2. This also means that the workload associated with the technical protocols for the single laboratory is comparable with the method comparison study of ISO 16140-2.

This document provides two strategies for the single-laboratory method validation of detection and quantification methods. The first strategy is based on a factorial approach while the second strategy uses the conventional approach derived from the protocols of ISO 16140-2. In addition, protocols for the determination of the in-house reproducibility for quantitative methods are described.

The advantages of using a factorial approach, over the conventional approach, are that it takes into account specific conditions that the laboratory encounters during routine testing and provides more information on the factors (technicians, culture media, etc.) that vary within the laboratory across relevant (food) items, while using fewer samples to assess the performance of the method. The factorial approach offers assessment of the precision of quantitative methods. It allows computation of reliable and representative single-laboratory method validation parameters such as in-house reproducibility standard deviation, LOD_{50} or RLOD values because it provides information on the variability of these values under different measurement conditions. The factorial approach requires fewer test results in order to obtain similar or higher levels of reliability compared to the conventional approach.

Microbiology of the food chain — Method validation —

Part 4: Protocol for method validation in a single laboratory

1 Scope

This document specifies the general principles and the technical protocols for single-laboratory validation of methods for microbiology in the food chain. The protocols in this document only validate the method for the laboratory conducting the study.

This document is applicable to single-laboratory validation of:

— methods used in the analysis (detection or quantification) of microorganisms in:
  — products intended for human consumption;
  — products intended for animal feeding;
  — environmental samples in the area of food and feed production, handling;
  — samples from the primary production stage;
— methods for the confirmation or typing of microorganisms. This validation will replace only the confirmation or typing procedure of a specified method (see Annex G).

This document is, in particular, applicable to bacteria and fungi. Some clauses can be applicable to other (micro)organisms or their metabolites, to be determined on a case-by-case basis.

Single-laboratory validation is required if an interlaboratory validation in accordance with ISO 16140-2 is not appropriate. Possible applications are:

— validation of an in-house method;
— method evaluation study in the validation process of a reference method in accordance with ISO 17468;
— extension of the scope of an ISO 16140-2 validated method, e.g. category extension or test portion size;
— modifications of existing methods.

Single-laboratory validation is the second step in the standardization of a reference method (see ISO 17468). It is only applicable to methods that are fully specified with regard to all relevant parameters (including tolerances on temperatures and specifications on culture media) and that have already been optimized.

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 6887 (all parts), *Microbiology of the food chain — Preparation of test samples, initial suspension and decimal dilutions for microbiological examination*

ISO 7218, *Microbiology of food and animal feeding stuffs — General requirements and guidance for microbiological examinations*

ISO 16140-1:2016, *Microbiology of the food chain — Method validation — Part 1: Vocabulary*

ISO 16140-2:2016, *Microbiology of the food chain — Method validation — Part 2: Protocol for the validation of alternative (proprietary) methods against a reference method*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 16140-1 and the following apply.

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

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

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

3.1
block
group of *settings* (3.12) that are conducted in parallel or in a short time interval, and that are used for the same samples

EXAMPLE Block = settings conducted in parallel =

technician “a” + culture medium “b” + temperature “a” + incubation condition “a”

and

technician “b” + culture medium “a” + temperature “b” + incubation condition “b”.

Note 1 to entry: This definition is based on how ISO 3534-3:2013, 3.1.25, defines “block”. In ISO 3534-3:2013, 3.1.25, the definition is more general as it is defining a block as a set of experimental units that are homogenous in some sense. The statistical meaning is the same.

3.2
factor
qualitative or quantitative parameter within the method that can be varied at two or more levels within the limits of the specified method

EXAMPLE Technician.

Note 1 to entry: In this document, only those factors that are in line with the protocol of the method are considered.

3.3
factor level
value of the *factors* (3.2) within the experimental design

EXAMPLE Technician “a”, technician “b”, etc.

Note 1 to entry: In this document, each factor is varied at two factor levels: “a” and “b”.

Note 2 to entry: This definition is based on how ISO 3534-3:2013, 3.1.12, defines “factor level”. In ISO 3534-3:2013, 3.1.12, the definition is more general, but the statistical meaning is the same.

3.4
in-house repeatability
measurement precision under a set of in-house repeatability conditions in a specific laboratory

Note 1 to entry: In-house repeatability conditions include the same measurement procedure, same technicians, same measuring system, same operating conditions, same location and replicate measurements on the same or similar objects over a short period of time in a particular laboratory.

3.5
in-house reproducibility
measurement precision under a set of in-house reproducibility conditions in a specific laboratory

Note 1 to entry: In-house reproducibility conditions include different technicians, different operating conditions and replicate measurements on the same or similar objects over a longer period of time in a particular laboratory.

3.6
level of detection
LOD_x
<qualitative methods> measured analyte concentration, obtained by a given measurement procedure, for which the *probability of detection* (3.9) is *x*

EXAMPLE LOD_{50} is the level of detection for which 50 % of tests give a positive result.

Note 1 to entry: The term "level of detection" is used for qualitative methods in microbiology based on replicate analyses with three different contamination levels of the target analyte in a tested matrix. The replicates are analysed, and the number of positive results is recorded (e.g. 20 %, 70 % and 100 %) respectively at each contamination level. These data are then used to determine the number of cells that would give 50 % positive using a generalized linear model (see ISO 16140-2). This differs from the procedure 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.

[SOURCE: ISO 16140-1:2016, 2.35, modified — Note 1 to entry has been slightly modified.]

3.7
limit of quantification
LOQ
limit of determination
<quantitative methods> lowest analyte concentration that can be quantified with an acceptable level of precision and trueness under the conditions of the test

[SOURCE: ISO 16140-1:2016, 2.36]

3.8
orthogonal design
factorial design, in which for every pair of *factors* (3.2), each combination of *factor levels* (3.3) occurs the same number of times across the possible factor levels

Note 1 to entry: This definition is based on how ISO 3534-3:2013, 3.1.31, defines "orthogonal array", but for "orthogonal design", a more general and more theoretical definition is used.

3.9
probability of detection
POD
proportion of positive analytical outcomes for a qualitative method for a given matrix at a given analyte level or concentration

Note 1 to entry: For qualitative methods, POD represents the probability of detection.

[SOURCE: ISO 16140-1:2016, 2.53, modified — Note 1 to entry has been added.]

3.10
relative level of detection
RLOD
level of detection (3.6) at $P = 0{,}50$ (LOD_{50}) of the alternative (proprietary) method divided by the level of detection at $P = 0{,}50$ (LOD_{50}) of the reference method

Note 1 to entry: For purposes of alternative-method acceptance, the derived RLOD is checked with the acceptability limit for conformity.

[SOURCE: ISO 16140-1:2016, 2.61]

3.11
single-laboratory method validation
in-house method validation
establishment of the performance characteristics of a method for the one particular laboratory in which the validation is conducted

3.12
setting
combination of *factor levels* (3.3)

EXAMPLE Technician "a" + culture medium "b" + temperature "a" + etc.

Note 1 to entry: These conditions can be described by the combination of levels of factors varied within the study.

4 General principles of the single-laboratory detection or quantification method validation

4.1 General

A single-laboratory detection or quantification method validation study is the first step in the framework of general method validation and is needed to assess the performance of the method across (food) categories, (food) types and (food) items. The second step in general method validation is an interlaboratory study to assess the performance of the method across laboratories.

A single-laboratory method validation study is used to demonstrate the performance of the method in the laboratory that conducted the study. The results are only valid for that particular laboratory.

NOTE Annex G gives the general principles for single-laboratory validation of alternative methods for microbiological confirmation and typing procedures.

This document describes two approaches for single-laboratory method validation:

— a factorial approach, with:
  — performance characteristics derived from ISO 16140-2;
  — an orthogonal, factorial study design (see ISO 3534-3);
  — more routine settings covered and fewer tests required than the conventional approach;
— a conventional approach, with:
  — performance characteristics derived from ISO 16140-2;
  — a stepwise procedure;
  — a study design derived from ISO 16140-2.

Validation protocols are dependent on whether the method is qualitative or quantitative, and on whether a factorial or a conventional approach is chosen. The factorial single-laboratory validation approach can only be used for a fully developed and optimized method. A conventional approach investigates the method for one specific setting (that is, one set of specific conditions under which the method is performed). The main differences in approach for the single-laboratory validation covered in this document are the number of various (food) items and the number of tests required to show that the method performs adequately. Validation of methods with, and without, a reference method is possible with the described protocols.

The scope of the validation protocol shall be determined at the start of the process, e.g. validation of in-house methods, the second step in the validation process in accordance with ISO 17468, extension of the scope of an ISO 16140-2 validated method, modification of existing methods.

For methods that include a PCR-based detection step, an assessment of the performance characteristics for the PCR-based detection step is described in ISO 22118. To ensure the reliable detection of the target organism in the samples tested, the relevant performance parameters of the PCR step should first be assessed (based on ISO 22118), before validation of the complete analytical procedure (e.g. following this document).

The selection of (food) categories and (food) types used in the validation study shall be conducted in accordance with ISO 16140-2:2016, 5.1.3.1. It is recommended that each (food) category relevant to the test method is also considered in the single-laboratory method validation study. Guidance on the selection of (food) categories and (food) types is given in ISO 16140-2:2016, Annex A.

The scope of the validation study, results (tables and calculations) of the different parts and the interpretation of the results, including discrepant results, shall be included in a validation study report.

4.2 Principles of the factorial approach

In a factorial approach, a systematic variation of factors is used to investigate the method performance under a defined range of conditions that are typically encountered in the routine application of the method. Typical factors are the technician or the sample storage, which can vary even within the same laboratory using a given procedure. By investigating the method in a variety of conditions concurrently, the factorial approach allows generalization of the validation to conditions commonly encountered in the laboratory and is not just limited to a single condition.

It is necessary to select four major factors that are expected to reflect the typical variation of conditions encountered in the routine application of the method. A risk analysis of the analytical workflow is recommended for the selection of the factors. Examples of factors are given in Annex A.

The systematic variation of factors ensures that their combined impact on general performance parameters, such as precision and sensitivity, can be derived. This is in contrast to a factorial robustness study, in which the central aim is the detection of specific significant method parameters, so that the performance of the method can be optimized. Compatibility between different factor levels and the impact on precision of non-significant effects are not taken into account in such a study.

Compared to the conventional approach as described in ISO 16140-2, the factorial approach requires a smaller number of (food) items and a smaller number of tests, while allowing for a reliable determination of validation parameters.

4.3 Principles of the conventional approach

The conventional approach principally follows ISO 16140-2. It is conducted in several steps and does not vary factors (see Table 2). The conventional approach requires more (food) items and test portions to be tested than the factorial approach.

Table 2 — Number of tests required for a method validation per (food) category by the factorial and conventional approach

<table>
<tr><th></th><th>Factorial approach</th><th>Conventional approach</th></tr>
<tr><td>Qualitative method against a reference method</td><td>
<table>
<tr><th></th><th>A</th><th>R</th></tr>
<tr><td>Factorial study (sensitivity + RLOD)</td><td>78</td><td>78</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td><td>0</td></tr>
<tr><td>Total number of tests</td><td colspan="2">236</td></tr>
</table>
(see 5.1.1)
</td><td>
<table>
<tr><th></th><th>A</th><th>R</th></tr>
<tr><td>Sensitivity study</td><td>60</td><td>60</td></tr>
<tr><td>RLOD study</td><td>30</td><td>30</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td><td>0</td></tr>
<tr><td>Total number of tests</td><td colspan="2">260</td></tr>
</table>
(see 6.1.1)
</td></tr>
<tr><td>Qualitative method without a reference method</td><td>
<table>
<tr><td>Factorial study (sensitivity + LOD_{50})</td><td>256</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td></tr>
<tr><td>Total number of tests</td><td>336</td></tr>
</table>
(see 5.1.2)
</td><td>
<table>
<tr><td>Specificity</td><td>20</td></tr>
<tr><td>LOD_{50} study (LOD_{50} + sensitivity)</td><td>360</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td></tr>
<tr><td>Total number of tests</td><td>460</td></tr>
</table>
(see 6.1.2)
</td></tr>
<tr><td>Quantitative method against a reference method</td><td>
<table>
<tr><th></th><th>A</th><th>R</th></tr>
<tr><td>Factorial study (relative trueness + accuracy profile + in-house precision)</td><td>48</td><td>48</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td><td>0</td></tr>
<tr><td>Total number of tests</td><td colspan="2">176</td></tr>
</table>
(see 5.2.1)
</td><td>
<table>
<tr><th></th><th>A</th><th>R</th></tr>
<tr><td>Relative trueness study</td><td>15</td><td>15</td></tr>
<tr><td>Accuracy profile study</td><td>30</td><td>30</td></tr>
<tr><td>In-house precision study</td><td>40</td><td>5</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td><td>0</td></tr>
<tr><td>Total number of tests</td><td colspan="2">215</td></tr>
</table>
(without LOQ study) (see 6.2.1)
</td></tr>
<tr><td>Quantitative method without a reference method</td><td>
<table>
<tr><td>Factorial study (relative trueness + accuracy profile + in-house precision)</td><td>48</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td></tr>
<tr><td>Total number of tests</td><td>128</td></tr>
</table>
(see 5.2.2)
</td><td>
<table>
<tr><td>Relative trueness study</td><td>15</td></tr>
<tr><td>Accuracy profile study</td><td>30</td></tr>
<tr><td>In-house precision study</td><td>40</td></tr>
<tr><td>Inclusivity/exclusivity study[a]</td><td>80</td></tr>
<tr><td>Total number of tests</td><td>165</td></tr>
</table>
(without LOQ study) (see 6.2.2)
</td></tr>
<tr><td colspan="3">Key
A: number of tests of the alternative method
R: number of tests of the reference method
[a] Inclusivity/exclusivity study requires 80 culture strains (130 for Salmonella) and is carried out only once for all approaches irrespective of the number of (food) categories.</td></tr>
</table>

5 Technical protocol for validation — Factorial approach

5.1 Qualitative methods

5.1.1 Single-laboratory method validation study against a reference method

5.1.1.1 General considerations

The factorial single-laboratory validation can only be used for a fully developed and optimized method. The validation study consists of two parts:

— a factorial, orthogonal comparison study (sensitivity and RLOD);

— an inclusivity/exclusivity study of the alternative method.

See Annex D for an elaborated example.

5.1.1.2 Factorial, orthogonal method comparison study

5.1.1.2.1 Selection of samples

The method comparison study compares the results obtained by the reference method and that of the alternative method. The study is conducted using naturally and/or artificially contaminated samples: usually, only artificially contaminated samples are used.

The requirements are as follows.

— Twelve different (food) items shall be selected for each (food) category: three (food) types per (food) category shall be selected and four (food) items shall be selected for each (food) type. (Food) items should be representative for the respective (food) type.

— The selection of (food) items shall take into account: background microbiota and food-processing factors, such as heat, pH, freezing, smoking, drying (low a_w); matrix conditions, such as pH value, a_w value, aerobic/anaerobic; special sample preparation requirements, such as high fat content or presence of inhibitors, in accordance with the ISO 6887 series.

— Each (food) item shall be contaminated at a minimum of two levels, consisting of at least the following.

  — A low (fractional) level L_1: the low level should have fractional recovery by the reference method (fractional recovery at the low level should be between 25 % and 75 % of the number of test portions tested). Ideally, the low level should be close to the theoretical detection level of 0,7 cfu/test portion (e.g. 0,5 cfu/test portion to 0,9 cfu/test portion).

  — A high level L_2: at the high level (e.g. 5 cfu/test portion to 10 cfu/test portion), 100 % positive results are expected.

— The four (food) items from each (food) type are allocated at random to four different blocks. Each block shall contain three (food) items, each at two contamination levels, from three (food) types.

— Use a different strain per block and/or the same strain subjected to different stress factors [e.g. temperature abuse, acid treatment or chlorination, depending on their relevance for the (food) type]. Where it is not possible to use different strains for each block, the laboratory needs to provide an explanation.

— Points to be considered when selecting strains are provided in ISO 16140-2:2016, Annex E.

— Six (food) items out of the twelve different (food) items shall be tested at zero level L_0 (blank, i.e. no target organism in the test portions).

— The size of the test portion shall be standardized for each study and should be the same for both methods, if possible. If the size of the test portion allowed by the alternative method is different from the reference method, contamination levels of the target microorganisms have to be adjusted accordingly so that the final contamination level (cfu/test portion) is the same for both.

— General protocols for artificial contamination of samples are provided in ISO 16140-2:2016, Annex C.

— Artificial contamination of samples shall be finalized before starting the analyses.

5.1.1.2.2 Selection of method factors

Decisions on the most suitable factors for the particular study should be based on expert knowledge. For example, optimal conditions are specified in each method (e.g. incubation temperature and duration at 37 °C and 24 h) and these will give the best results. However, ranges around these, which provide still acceptable conditions (e.g. for 24 h ± 1 h), are permitted and the study design should test the extremes of these (e.g. incubate the samples for 23 h or 25 h). Acceptable operating conditions for equipment are described in ISO 7218.

Relevant method factors that are more difficult to control (e.g. technicians, culture media and incubation temperature) shall be selected and varied systematically to enable assessment of the accuracy under routine conditions in the specific laboratory. The choice of these factors and factor levels is crucial to the reliability of the validation result. These shall reflect the variation within the single laboratory under routine conditions and should cover the most relevant aspects of the method, such as sample preparation, sample storage, laboratory technician, laboratory equipment or background microbiota. Other influences, such as atmosphere and stress conditions, can also be taken into account.

Four relevant factors shall be varied simultaneously, each on two levels.

For methods for culturable microorganisms, the factors and factor levels shown below are to be taken into consideration. “Technicians” and “culture media” have the greatest impact and shall be included in all studies as follows.

— Technicians: Tests shall be independently conducted in the single laboratory by two technicians.

— Culture media: Use culture media from two different manufacturers, if available, or two different batches of culture media (lots), or pre-prepared versus prepared from dehydrated media. The choices depend on the normal conditions of use of media in the laboratory. For example, two different batches can be used if only one product is used in the laboratory, even if the product is available from different manufacturers.

Two other factors shall be taken into account. A non-comprehensive list of grouped potential factors is provided in Annex A. If possible, one factor from two of the most relevant groups shall be selected.

Factors are studied simultaneously using the study design described in 5.1.1.2.3.

5.1.1.2.3 Experimental design

The twelve (food) items from three (food) types shall be allocated to four blocks: that is, each block shall include one (food) item from each (food) type. Each (food) item, contaminated at fractional and high level, is analysed under two different settings (factor level combinations). Each setting is a combination of levels of four factors, e.g. (food) items 1 to 3 are analysed in setting 1: technician “a”, background microbiota “b”, culture medium “a”, incubation condition “a” (see Table 3). In addition, six selected (food) items shall be tested at zero level L_0 (blank).

Tests shall be performed as follows:

— the zero level L_0 (blank) shall be tested using 1 replicate for each of 6 selected (food) items representing the 3 (food) types (6 tests);

NOTE 1 The zero level L_0 samples are tested to demonstrate that there are no false positive results [cross reactivity e.g. with (food) matrix]. Therefore, the choice of the settings used for testing the L_0 samples is not important, i.e. any setting can be used. All food items used (all 12 and not only the 6 used for zero level) are examined so that these do not contain the target organism.

— the fractional level L_1 shall be tested using 2 settings × 2 replicates for each of the 12 (food) items in accordance with the factorial design presented in Table 3 (48 tests);

— the high level L_2 shall be tested using 2 settings × 1 replicate for each of the 12 (food) items in accordance with the factorial design presented in Table 3 (24 tests).

Therefore, a total of 78 single tests per (food) category examined are to be conducted per test method. That is, 156 single tests are to be conducted for the two (reference and alternative) methods.

NOTE 2 The second setting within each block corresponds to the first setting with all factor levels switched around. The second setting therefore reflects the most extreme deviation from the first setting.

NOTE 3 The factorial design with blocks allows not only detection of various effects of factors, but also of interactions between factors.

NOTE 4 As an illustration, consider the case where all results were in the expected range, except the results of setting 3 with factor level combination a-a-a-b and setting 8 with factor level combination b-a-a-a. This suggests that background microbiota "a" causes an undesired effect for culture medium "a", because this is what the two settings have in common. In this case, the laboratory would perform a root cause analysis in order to provide an explanation for the observed results.

Table 3 — Study design for a factorial, orthogonal study for qualitative methods against a reference method; per (food) category

12 (food) items from 3 (food) types in random order, each (food) item with known contamination levels	6 out of the 12 (food) items at zero level	6 out of the 12 (food) items at any setting							
	The 12 (food) items at fractional level	1 to 3		4 to 6		7 to 9		10 to 12	
	The 12 (food) items at high level	1 to 3		4 to 6		7 to 9		10 to 12	
		Block 1		Block 2		Block 3		Block 4	
Setting		1	2	3	4	5	6	7	8
Factor 1	technician	a	b	a	b	a	b	a	b
Factor 2	culture medium	b	a	a	b	a	b	b	a
Factor 3	e.g. background microbiota (storage)	a	b	a	b	b	a	b	a
Factor 4	e.g. incubation condition	a	b	b	a	a	b	b	a

NOTE 1 Allocation of the 12 (food) items to 4 blocks, each with 2 different factor-level combinations (settings). The 12 (food) items are contaminated at 2 levels: fractional (2 replicates) and high (1 replicate).

NOTE 2 Number of tests: 12 (food) items × ((2+1) × replicates) × 2 settings × 2 methods = 144 tests.

NOTE 3 In addition to the 144 tests: 6 (food) items × 1 replicate × 2 methods = 12 tests for zero level L_0 (blank).

NOTE 4 Total number of tests = 144 + 12 = 156.

Figure 3 provides an overview of the design for technician "a".

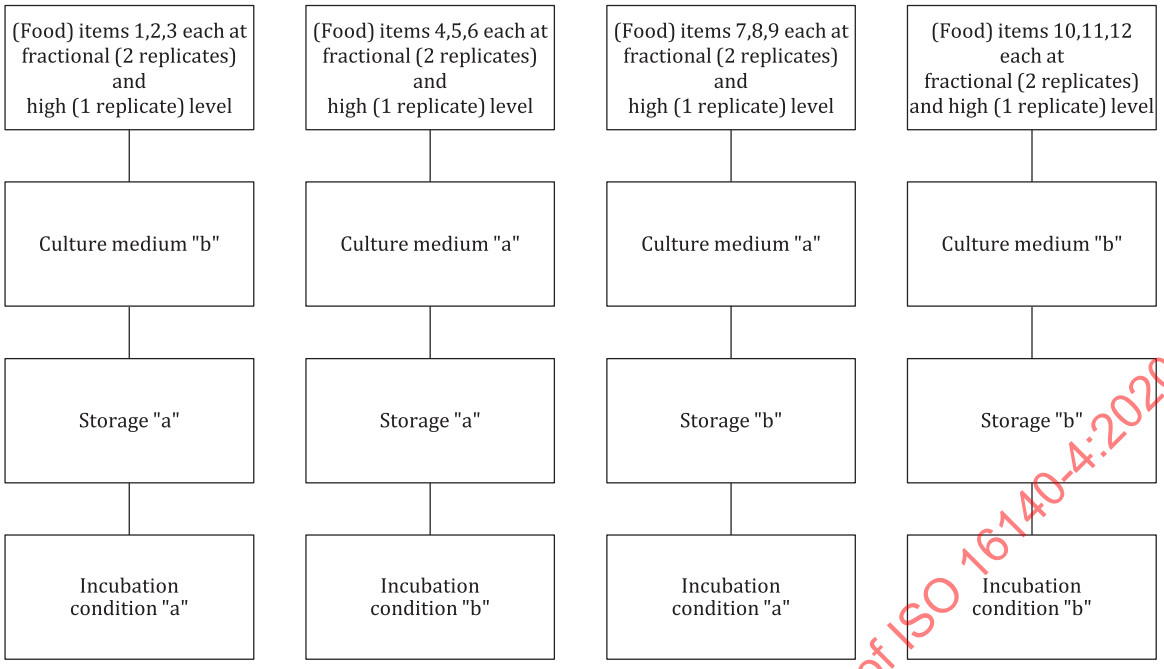


Figure 3 — Design for technician "a"

Confirmation of the alternative method (end) result is needed as described in ISO 16140-2:2016, 5.1.3.3, to avoid false-positive results and confirm true positive results.

An example of a factorial single-laboratory method validation study for a qualitative method against a reference method is provided in Annex D.

5.1.1.3 Calculation and interpretation for sensitivity

Calculations based on the results of the factorial comparison study at fractional level L_1 shall be conducted in accordance with ISO 16140-2:2016, 5.1.3.4. All calculations shall be performed:

— separately, for each (food) category [all (food) types];

— separately, for each (food) type.

The results obtained can be summarized as presented in Table 4.

The results and calculations obtained for (food) categories are used to evaluate against the acceptability limits (see Clause 7). The results and calculations obtained for (food) types can only be used as indicative values and can also be used in a root cause analysis.

To see factorial effects, calculations can also be conducted for specific factor levels, e.g. for all results from culture medium "a", across (food) types and (food) categories.

To see factorial interaction effects, examine:

a) whether the results for 2 of the 8 settings are not in the expected range;

b) which factor levels these two settings have in common.

For example, settings 2 and 5 have factor levels of culture medium (level "a") and background microbiota (level "b") in common, i.e. unexpected results in settings 2 and 5 suggest that there is an interaction effect between culture medium and background microbiota.

Table 4 — Summary of results obtained with the reference and alternative method at fractional level L_1 for each (food) category and (food) type, based on the study design described in Table 3

(Food) category	(Food) types	PA	NA	ND	PD	FP	N[a]	$SE_{(alt)}$ %	$SE_{(ref)}$ %	RT[b] %	FPR[c] %
Category 1	All (food) types										
Category 1	(Food) type A										
Category 1	(Food) type B										
Category 1	(Food) type C										
Category 1	Technician “a” [all (food) types]										
Category 1	Technician “b” [all (food) types]										
Category 1	Culture medium “a” [all (food) types]										
Category 1	Culture medium “b” [all (food) types]										
Category 1	Background microbiota “a” [all (food) types]										
Category 1	Background microbiota “b” [all (food) types]										
Category 1	Incubation condition “a” [all (food) types]										
Category 1	Incubation condition “b” [all (food) types]										
Category 2	All (food) types										
Category 2	(Food) type A										
Category 2	(Food) type B										
Category 2	(Food) type C										
Category 2	… (separately for each factor level)										
Category x	All (food) types										
Category x	(Food) type A										
Category x	(Food) type B										
Category x	(Food) type C										
Category x	… (separately for each factor level)										

Key

PA: positive agreement, NA: negative agreement, ND: negative deviation, PD: positive deviation, FP: false positive, N: sum, $SE_{(alt)}$: sensitivity alternative method, $SE_{(ref)}$: sensitivity reference method, RT: relative trueness, FPR: false positive ratio

[a] N = NA + PA + PD + ND.

[b] RT = (PA + NA)/N × 100 %.

[c] FPR = FP/NA × 100 %.

NOTE Definitions and the derivation of the parameters can be found in ISO 16140-2:2016, 5.1.3.4.

Calculate the difference, (ND – PD), for both paired and unpaired studies and the sum, (ND + PD), for paired studies. The explanation for paired and unpaired studies can be found in ISO 16140-2:2016, 5.1.2. Check whether the difference and/or sum of PD and ND conform to the acceptability limit (AL) stated in Table 5. The AL is not met when the observed value is higher than the AL. The interpretation of results shall be done per (food) category and for all (food) categories used in the validation study. See Annex D for an elaborated example.

Table 5 — Acceptability limit (AL) parameters and values for a paired and unpaired study design in relation to the number of categories used

Number of categories	Paired study		Unpaired study
	(ND[a] – PD[b])	(ND + PD)	(ND – PD)
1	3	6	3
2	4	8	4
3	5	10	5
4	5	12	5
5	5	14	5
6	6	16	6
7	6	18	7
8	6	20	7

[a] ND = number of samples with negative deviation results at fractional level L_1.

[b] PD = number of samples with positive deviation results at fractional level L_1.

NOTE Acceptability limits (AL) are based on data and consensus expert opinion. The AL are not based on statistical analysis of the data.

5.1.1.4 Calculation and interpretation of the RLOD

Calculations based on all results of the factorial comparison study per (food) category shall be conducted in accordance with ISO 16140-2:2016, 5.1.4.2.

Effects of the individual factors should be analysed based on the RLOD values calculated for each factor level.

NOTE An Excel®-based program[1)] is available for performing the RLOD calculations from https://standards.iso.org/iso/16140/-4/ed-1/en.

For a specific factor of interest, denote by:

— $Y^{(a)}$ the $\log_{10}$ value of the RLOD using all test results of the four settings in which the factor is set to level "a";

— $Y^{(b)}$ the $\log_{10}$ value of the RLOD using all test results of the other four settings in which the factor is set to level "b".

Calculate the difference of these $\log_{10}$ RLOD values as shown by Formula (1):

$$d = Y^{(b)} - Y^{(a)} = \log_{10}(\text{RLOD}(\text{level "b"})) - \log_{10}(\text{RLOD}(\text{level "a"})) \quad (1)$$

If this difference d is smaller than −0,6 or larger than +0,6, the factor is considered to have a substantial influence on the RLOD. In this case, it can be concluded that the LOD_{50} of the two factor levels differ by more than 4:1. This could be an indication of a problem with the alternative or the reference method. Perform a root cause analysis in order to provide an explanation for the out of range results. Reject the alternative method if it cannot be excluded that there is a problem with this method.

Table D.4 provides an elaborated example. Further analysis techniques can be found in Reference [12].

5.1.1.5 Inclusivity/exclusivity study

In the inclusivity/exclusivity study, both target strains and non-target strains are tested with the alternative method. The study, including the selection and number of strains, shall be conducted in

1) Excel® is the trade name of a product supplied by Microsoft and is an example of a suitable product available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of this product.

accordance with ISO 16140-2:2016, 5.1.5. If the validation is a matrix [(food) item or type or category] extension of an already validated method, no additional inclusivity/exclusivity study is required. If the enrichment procedure is replaced with another selective enrichment, an inclusivity/exclusivity study is required.

5.1.2 Single-laboratory method validation study without a reference method

5.1.2.1 General considerations

Without a reference method, the validation of the method is slightly varied from 5.1.1. Instead of a comparison study based on the RLOD, all calculations are based solely on the LOD_{50} values corresponding to each setting of the factorial design. The study is conducted using samples with known contamination levels and the inoculum should be enumerated using a non-selective medium. Enumeration shall be performed as described in ISO 7218.

5.1.2.2 Factorial study

Sixteen (food) items, from at least three (food) types, shall be selected for each (food) category. (Food) items shall be selected in accordance with the criteria described in 5.1.1.2. Each (food) item shall be artificially contaminated at two different levels: a low (fractional) level and a high level of contamination (ISO 16140-2:2016, 5.1.4.1). An uninoculated sample (blank) shall also be prepared for each (food) item. Tests shall be performed in replicate, with two replicates for the blank and high level, and four replicates for the fractional level.

Each (food) item is analysed under two different settings (= combination of levels for four factors). Thus, each of sixteen (food) items is analysed in two settings with two replicates at both a zero and a high level, and with four replicates at a low (fractional) level. In total 256 [16 (food) items × (2 + 4 + 2) replicates × 2 settings] tests are performed.

The experimental design of 5.1.1.2 is followed for selection of factors. Table 6 shows the study design for the factorial, orthogonal study for qualitative methods without a reference method for one (food) category.

Table 6 — Study design for a factorial, orthogonal study for qualitative methods without a reference method; per (food) category

16 (food) items from at least 3 (food) types from 1 (food) category, in random order, each (food) item with known contamination levels: zero, fractional and high	The 16 (food) items at zero level	1 to 4		5 to 8		9 to 12		13 to 16	
	The 16 (food) items at fractional level	1 to 4		5 to 8		9 to 12		13 to 16	
	The 16 (food) items at high level	1 to 4		5 to 8		9 to 12		13 to 16	
		Block 1		Block 2		Block 3		Block 4	
Setting		1	2	3	4	5	6	7	8
Factor 1	technician	a	b	a	b	a	b	a	b
Factor 2	culture medium	b	a	a	b	a	b	b	a
Factor 3	e.g. background microbiota (storage)	a	b	a	b	b	a	b	a
Factor 4	e.g. incubation condition	a	b	b	a	a	b	b	a

NOTE 1 Allocation of the 16 (food) items to 4 blocks, each with 2 different factor-level combinations (settings). The 16 (food) items have known contamination levels: zero (2 replicates), fractional (4 replicates) and high (2 replicates).

NOTE 2 Number of tests: 16 (food) items × (2 + 4 + 2) replicates × 2 settings = 256 tests.

The validation study should be performed using four strains per (food) category. Use a different strain per block and/or the same strain subjected to different stress factors [e.g. temperature abuse, acid treatment or chlorination, depending on their relevance for the (food) type].

When testing different categories also different strains per category should be used ideally. However, this will depend on the availability of strains.

Where it is not possible to use different strains, for each block or the entire study, the laboratory should provide an explanation.

Points to be considered when selecting strains are provided in ISO 16140-2:2016, Annex E.

5.1.2.3 Calculation and interpretation for sensitivity

Calculation of the results of the factorial study at zero level and at high level shall be based on the principles outlined in ISO 16140-2:2016, 5.1.3.4, for an unpaired study. Test results at low (fractional) level are not used. All reference results at zero level are treated as negative results. At high level, they are treated as positive results.

NA represents the number of negative test results and PD the number of positive test results of the method at zero level. ND represents the number of negative test results and PA the number of positive test results of the method at high level.

Acceptability limits (ALs) per (food) category are as follows: ND = 3, PD = 1. If ND is larger than 3, or PD is larger than 1, the method cannot be validated.

5.1.2.4 Calculation and interpretation of LOD_{50}

The LOD_{50} can be calculated for each setting using the complementary-log-log (CLL) approach for the probability of detection (POD) adapted from ISO 16140-2. The LOD_{50} across settings can be obtained as 10^k where k denotes the mean of $\log_{10}$ LOD_{50} values across settings (i.e. as the geometric mean of the LOD_{50} values). An example is provided in Annex E. A more comprehensive approach for the calculation of the LOD_{50} that includes the calculation of in-house reproducibility standard deviation can be found in Reference [12].

NOTE 1 The LOD_{50} can be calculated using "R". "R" is a programming language and free software environment for statistical computing and graphics that is supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians. It can be downloaded from https://cran.r-project.org. The R code for the computation of LOD_{50} is as follows:

```
fit=glm(Result~1+offset(log(Concentration))+Item,
    family=binomial(link="cloglog"), data=TEST.RESULTS.DATAFRAME)
Item1 = summary(fit)$coefficients[1]
Item2 = summary(fit)$coefficients[1]+summary(fit)$coefficients[2]
(…)
ItemN = summary(fit)$coefficients[1]+summary(fit)$coefficients[N]
MeanAcrossItems = mean(c(Item1, Item2,…, ItemN))
LOD50 = -log(0,5)/exp(MeanAcrossItems)
```

NOTE 2 An Excel®-based program[2)] is available for performing the LOD_{50} calculations from https://standards.iso.org/iso/16140/-4/ed-1/en.

2) Excel® is the trade name of a product supplied by Microsoft and is an example of a suitable product available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of this product.

In addition, effects of the individual factors (e.g. technician) and of interactions between factors (e.g. culture medium and incubation conditions) can be calculated from the factorial design based on the LOD_{50} values calculated for each setting. For a specific factor of interest, denote by:

— $Y^{(a)}$ the $\log_{10}$ value of the LOD_{50} value using all test results of the four settings in which the factor is set to level "a";

— $Y^{(b)}$ the $\log_{10}$ value of the LOD_{50} values using all test results of the other four settings in which the factor is set to level "b".

Calculate the difference of these $\log_{10}$ LOD_{50} values as shown by Formula (2):

$$d = Y^{(b)} - Y^{(a)} \quad (2)$$

If this difference d is smaller than −0,6 or larger than +0,6, the factor is considered to have a substantial influence on the LOD_{50}. In this case, it can be concluded that the LOD_{50} of the two factor levels differ by more than 4:1. This could be an indication of an undesired effect with the alternative method. Perform a root cause analysis in order to provide an explanation for the large differences in results. Examine whether the undesired effect is related to certain factor level combinations. Reject the alternative method if it cannot be excluded that there is a problem with this method.

5.1.2.5 Inclusivity/exclusivity study

Inclusivity/exclusivity testing shall be conducted in accordance with 5.1.1.5.

5.2 Quantitative methods

5.2.1 Single-laboratory method validation study against a reference method

5.2.1.1 General considerations

The factorial single-laboratory validation can only be used for a fully developed and optimized method. The validation study consists of two parts:

— a factorial, orthogonal comparison study (relative trueness and accuracy profile and in-house precision);

— an inclusivity/exclusivity study of the alternative method.

See Annex C for an elaborated example.

5.2.1.2 Selection of samples

The method comparison study compares the results obtained by the reference method with that of the alternative method. The study is conducted using naturally and/or artificially contaminated samples, selected as follows.

— For each (food) category to be tested, twelve (food) items from three or four (food) types shall be selected. For each (food) type, at least three (food) items shall be selected. For example, if there are three (food) types, four (food) items could be selected for each (food) type, but it is also possible to select six (food) items for the first (food) type and three (food) items for each of the two remaining (food) types.

— Selection shall take into account background microbiota and food-processing factors, such as heat, pH, freezing, smoking and drying (a_w), and matrix conditions, such as pH value, a_w value, aerobic/anaerobic and resuscitation procedures.

— The size of the test portion shall be standardized for each study.

— (Food) items may be naturally or artificially contaminated. In the case of artificial contamination, use a different strain and/or background microbiota and stress factors, for each (food) item, where possible, e.g. temperature abuse, acid treatment or chlorination, depending on relevance for the respective (food) type. Where it is not possible to use different strains per (food) item, the laboratory needs to provide an explanation.

— Points to be considered when selecting strains are provided in ISO 16140-2:2016, Annex E.

— The twelve (food) items shall be (artificially or naturally) contaminated at three levels: four (food) items shall be at low contamination level L_1 (denoted by numbers 1, 4, 7, 10), four at medium contamination level L_2 (2, 5, 8, 11), and four at high contamination level L_3 (3, 6, 9, 12), see Table 7.

— General protocols for the artificial contamination of samples are provided in ISO 16140-2:2016, Annex C.

5.2.1.3 Selection of method factors

The selection of method factors shall be performed as described in 5.1.1.2.2. Factors are studied simultaneously using the study design described in 5.2.1.4.

5.2.1.4 Experimental design

— For each method and (food) category: twelve (food) items (four at low, four at medium and four at high contamination level) shall be analysed in two settings. Two replicates, i.e. two independent dilutions of the (food) item, shall be conducted in each setting. Therefore, 12 (food) items × 2 settings × 2 replicates = 48 single tests are to be conducted. A total of 96 (48 for the reference method and 48 for the alternative method) single tests are performed per (food) category.

— Each setting is a combination of levels of four method factors. Eight different settings (1 to 8) shall be considered.

Table 7 shows the allocation of samples to the different settings.

Table 7 — Study design for a factorial validation study for a quantitative method with a reference method; per (food) category

12 (food) items from at least 3 (food) types per (food) category, in random order	4 (food) items at low level	1		4		7		10	
	4 (food) items at medium level	2		5		8		11	
	4 (food) items at high level	3		6		9		12	
		Block 1		Block 2		Block 3		Block 4	
Setting		1	2	3	4	5	6	7	8
Factor 1	technician	a	b	a	b	a	b	a	b
Factor 2	culture medium	b	a	a	b	a	b	b	a
Factor 3	e.g. incubation duration	a	b	a	b	b	a	b	a
Factor 4	e.g. time of reading	a	b	b	a	a	b	b	a
NOTE Number of tests per method: 12 (food) items × 2 settings × 2 replicates = 48 tests.									

5.2.1.5 Relative trueness

Calculate the average for the four (2 settings × 2 replicates) test results for each (food) item and use these averages for the calculation of relative trueness across all settings. Calculation and interpretation of relative trueness shall be conducted in accordance with ISO 16140-2:2016, 6.1.2.3.

In addition, determine the relative trueness separately for each level of the four factors in accordance with the allocation of settings given in Table 7: e.g. technician "a" (settings 1, 3, 5, 7), technician "b" (settings 2, 4, 6, 8), culture medium "a" (settings 2, 3, 5, 8), culture medium "b" (settings 1, 4, 6, 7), incubation duration "a" (settings 1, 3, 6, 8), incubation duration "b" (settings 2, 4, 5, 7), time of reading "a" (settings 1, 4, 5, 8), time of reading "b" (settings 2, 3, 6, 7).

Calculate the differences in bias:

a) between the two technicians;

b) between the two culture media;

c) between the two incubation durations;

d) between the two different times of reading.

NOTE Further tests can be performed if certain factor combinations yield unexpected results. The factorial design allows further interrogation of performance parameters, if there is a particularly significant result for specific combination of factors, e.g. culture medium in combination with a specific incubation setting.

5.2.1.6 Accuracy profile

Calculation and interpretation of the accuracy profile shall be conducted in accordance with ISO 16140-2:2016, 6.1.3.3. Treat the four test results (2 settings × 2 replicates) from each method for each (food) item as four replicates.

NOTE 1 Further assessments of the accuracy profile can be carried out by analysis of factorial effects.

NOTE 2 An Excel®-based program[3)] is available for performing the accuracy profile calculations from https://standards.iso.org/iso/16140/-4/ed-1/en.

5.2.1.7 In-house precision (in-house repeatability and in-house reproducibility)

The following notation is used: i refers to the level ($1 \le i \le 3$); j refers to the setting ($1 \le j \le 8$); k refers to the replicate ($1 \le k \le 2$). Note the data as follows:

— y_{ijk}, is the $\log_{10}$ transformed test result on level i for setting j of replicate k using the alternative method;

— y_{p1} and y_{p2}, are the two mean values of the paired values y, with p denoting the respective sample p ($1 \le p \le 12$) of the two corresponding settings.

3) Excel® is the trade name of a product supplied by Microsoft and is an example of a suitable product available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of this product.

The in-house repeatability standard deviation across (food) types is computed as shown by Formula (3):

$$s_{r,\text{in-house}} = \sqrt{\frac{1}{48}\sum_{i=1}^{3}\sum_{j=1}^{8}\left(y_{ij1} - y_{ij2}\right)^2} \tag{3}$$

The in-house reproducibility standard deviation across (food) types is computed as shown by Formula (4):

$$s_{R,\text{in-house}} = \sqrt{s^2_{L,\text{in-house}} + s^2_{r,\text{in-house}}} \tag{4}$$

where

$$s_{L,\text{in-house}} = \sqrt{\frac{1}{24}\sum_{p=1}^{12}\left(y_{p1} - y_{p2}\right)^2 - \frac{1}{2}s^2_{r,\text{in-house}}}$$

If the formulae result in the computation of the square root of a negative number, the result = 0.

In-house repeatability standard deviation and in-house reproducibility standard deviation are used to describe precision of the alternative method under in-house repeatability conditions and in-house reproducibility conditions.

NOTE 1 The variance components can be calculated more efficiently, with better precision, using restricted maximum likelihood (REML) estimates[13].

NOTE 2 Evaluation of the calculated in-house reproducibility standard deviation can be conducted using the criterion proposed by Reference [11]: target value ±0,25 $\log_{10}$ cfu/g. Since accuracy comprises precision, no additional evaluation of precision is required in this document.

5.2.1.8 Inclusivity/exclusivity study

Inclusivity/exclusivity testing is required for enumeration methods designed for specific microorganisms. It is not required for general enumeration methods such as total plate count and yeast and mould methods.

In the inclusivity/exclusivity study, both target strains and non-target strains are tested with the alternative method. The study, including the selection and number of strains, shall be conducted in accordance with ISO 16140-2:2016, 6.1.5. If the validation is a matrix [(food) item or type or category] extension of an already validated method, no additional inclusivity/exclusivity study is required.

5.2.2 Single-laboratory method validation study without a reference method

5.2.2.1 General considerations

The general considerations of 5.2.1.1 apply. Without a reference method, however, the results of the alternative method shall be compared with known contamination levels of the samples, and the inoculum should be enumerated using a non-selective medium. Enumeration shall be performed as described in ISO 7218.

If contamination levels are not known, relative dilution levels of the inoculum can be used to assess dilution proportionality of the method. This is a minimum requirement of each quantitative method: results of different dilution levels should be inversely proportional to the dilution level. Dilution proportionality and in-house repeatability precision can be examined by using the protocol of 5.2.1, where the reference method is substituted with the alternative method with additional dilution steps (e.g. diluting all samples 1:10).

5.2.2.2 Selection of samples

Samples shall be selected in accordance with 5.2.1.2.

5.2.2.3 Selection of method factors

Method factors shall be selected in accordance with 5.2.1.3.

5.2.2.4 Experimental design

(Food) items shall have a known contamination level at three levels: a low level L_1, a medium level L_2 and a high level L_3. Relative contamination levels shall be achieved by using different dilutions of the inoculum (e.g. 1:10:100). The experimental design from 5.2.1.4 is used: four (food) items per contamination level are analysed in two settings with two replicates. The allocations of the twelve different (food) items to the eight settings in four blocks are given in Table 7. In total, eight settings with three levels and two replicates are tested, yielding 48 individual tests.

NOTE For some methods, it could be more appropriate to use the same block of (food) items for all contamination levels of the same setting. This can be accomplished by using four sample blocks and assigning blocks to settings in accordance with Table 6.

5.2.2.5 Relative trueness

Calculation and interpretation of the relative trueness shall be based on the protocol of 5.2.1.5, including the calculation of relative trueness for the eight described subsets of data. This protocol is modified by using the known artificial contamination levels instead of the test results of the reference method.

5.2.2.6 Accuracy profile

A modified version of the accuracy profile shall be calculated by following ISO 16140-2:2016, 6.1.3, and replacing the test results of the reference method with the known contamination level of the sample. In addition, the following adaptations shall be made.

— The calculation of the standard deviation of the reference method in Step 5 of ISO 16140-2:2016, 6.1.3.3, can be omitted, and the value replaced by 0 since the contamination levels are known and, hence, there is no variability.

— Step 9 of ISO 16140-2:2016, 6.1.3.3, does not apply because the standard deviation of the reference values cannot exceed the threshold of 0,125.

5.2.2.7 In-house precision (in-house repeatability and in-house reproducibility)

In-house repeatability standard deviation and in-house reproducibility standard deviation are used to describe precision of the alternative method under in-house repeatability conditions and in-house reproducibility conditions. They are calculated in accordance with 5.2.1.7.

5.2.2.8 Inclusivity/exclusivity study

Inclusivity/exclusivity testing shall be conducted in accordance with 5.2.1.8.

6 Technical protocol for validation — Conventional approach

6.1 Qualitative methods

6.1.1 Single-laboratory method validation study against a reference method

6.1.1.1 General

The conventional approach for the single-laboratory validation is based on the ISO 16140-2:2016, 5.1, method comparison study design.

6.1.1.2 Sensitivity study

The sensitivity study compares the results obtained by the reference method with that of the alternative method. It shall be conducted and interpreted in accordance with ISO 16140-2:2016, 5.1.3. For each (food) category being examined, a minimum of 60 samples shall be tested. At least 20 samples representative of each of the three (food) types shall be tested per (food) category. Fractional positive results by either the reference or alternative method (i.e. samples should not be all positive or all negative) shall be obtained for each (food) type tested. In the ideal situation, 10 samples (50 %) tested per (food) type should be positive. For each (food) category, at least 30 samples shall have a positive result by the reference and/or the alternative method.

6.1.1.3 RLOD study

The RLOD study is a comparative study of the LOD_{50} obtained by the reference method and that of the alternative method. It shall be conducted and interpreted in accordance with ISO 16140-2:2016, 5.1.4.

6.1.1.4 Inclusivity/exclusivity study

Inclusivity/exclusivity testing shall be conducted in accordance with 5.1.1.5.

6.1.2 Single-laboratory method validation study without a reference method

6.1.2.1 General

If no reference method is available, the validation of the alternative method proceeds in two steps. The specificity of the method is first established, before investigation of the LOD_{50} and sensitivity, and performance of the inclusivity/exclusivity study.

For the selection of (food) categories and (food) items, see ISO 16140-2:2016, 5.1.3.1.

6.1.2.2 Specificity

Twenty non-contaminated (with the target microorganism) (food) items per (food) category shall be used as blank samples. If more than one sample yields a positive result, the method cannot be validated.

6.1.2.3 LOD_{50} study

For each (food) category, twelve (food) items from three (food) types shall be selected. Each (food) item shall be analysed on a different day. For each (food) item and for each day, 30 test portions are artificially contaminated at different levels:

— 5 test portions without contamination (uninoculated) so that no positive results are expected;

— 20 test portions at a fractional level such that about 50 % positive results are expected;

— 5 test portions at a high level, such that 100 % positive results are expected.

Use a different strain per (food) item and/or the same strain subjected to different stress factors [e.g. temperature abuse, acid treatment or chlorination, depending on their relevance for the (food) type]. Where it is not possible to use different strains for each (food) item, the laboratory needs to provide an explanation. Points to be considered when selecting strains are provided in ISO 16140-2:2016, Annex E. General protocols for artificial contamination of samples are provided in ISO 16140-2:2016, Annex C. Each inoculum should be enumerated using a non-selective medium. Enumeration shall be performed as described in ISO 7218.

In total 360 tests are conducted per (food) category.

For each (food) item, the LOD_{50} shall be calculated by using the complementary-log-log (CLL) approach for the probability of detection (POD) adapted from ISO 16140-2.

NOTE 1 The LOD_{50} can be calculated using "R". "R" is a programming language and free software environment for statistical computing and graphics that is supported by the R Foundation for Statistical Computing. The R language is widely used amongst statisticians. It can be downloaded from https://cran.r-project.org. The R code for the computation of LOD_{50} is as follows:

```
fit=glm(Result~1+offset(log(Concentration))+Item,
    family=binomial(link="cloglog"), data=TEST.RESULTS.DATAFRAME)
Item1 = summary(fit)$coefficients[1]
Item2 = summary(fit)$coefficients[1]+summary(fit)$coefficients[2]
(...)
ItemN = summary(fit)$coefficients[1]+summary(fit)$coefficients[N]
MeanAcrossItems = mean(c(Item1, Item2,..., ItemN))
LOD50 = -log(0,5)/exp(MeanAcrossItems)
```

NOTE 2 An Excel®-based program[4] is available for performing the LOD_{50} calculations from https://standards.iso.org/iso/16140/-4/ed-1/en.

Annex B provides further information on the LOD_{50} study that allows the calculation of in-house reproducibility.

6.1.2.4 Calculation and interpretation for sensitivity

Calculation for sensitivity shall be conducted in accordance with ISO 16140-2:2016, 5.1.3.4. Use the test results of the LOD_{50} study at zero level and at high level. Test results at low (fractional) level are not used. All reference results at zero level are treated as negative results. At high level, they are treated as positive results.

NA represents the number of negative test results of the method at zero level and PD the number of positive test results at zero level. ND represents the number of negative test results of the method at high level and PA the number of positive test results at high level.

Acceptability limits (ALs) per (food) category are as follows: ND = 3, PD = 1. If ND is larger than 3, or PD is larger than 1, the method cannot be validated.

6.1.2.5 Inclusivity/exclusivity study

Inclusivity/exclusivity testing shall be conducted in accordance with 5.1.1.5.

6.2 Quantitative methods

6.2.1 Single-laboratory method validation study against a reference method

6.2.1.1 General

The conventional approach for the single-laboratory validation is based on the ISO 16140-2:2016, 6.1, method comparison study design.

4) Excel® is the trade name of a product supplied by Microsoft and is an example of a suitable product available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of this product.

6.2.1.2 Relative trueness study

The relative trueness study is a comparative study between the results obtained by the reference method and that of the alternative method. It shall be conducted in accordance with ISO 16140-2:2016, 6.1.2.

6.2.1.3 Accuracy profile study

The accuracy profile study compares the results obtained by the reference method to that of the alternative method. It shall be conducted in accordance with ISO 16140-2:2016, 6.1.3.

6.2.1.4 Limit of quantification study

For some alternative methods, it is appropriate to determine the limit of quantification (LOQ). The LOQ is only relevant, and shall be determined, when the measurement principle of the alternative method is not based on the visual observation of the target microorganism. The LOQ study, if applicable, shall be conducted in accordance with ISO 16140-2:2016, 6.1.4.

6.2.1.5 In-house precision study

In-house repeatability standard deviation and in-house reproducibility standard deviation are important performance characteristics of any measurement method. They describe precision of the alternative method under in-house repeatability conditions and in-house reproducibility conditions.

One (food) item is selected per (food) category for the determination of the in-house repeatability standard deviation and the in-house reproducibility standard deviation.

If the artificially contaminated (food) item is sufficiently stable, microbiologically, tests are conducted on eight different days by three or more technicians. No technician shall conduct tests on more than three days. However, if only two technicians are available, then each technician shall conduct tests on four days. If only one technician is available, this technician shall conduct all tests, but the method is then validated only for this technician.

On each day, one of the technicians performs tests on five replicates using the alternative method, yielding a total of 8 × 5 = 40 tests. In addition, five replicate tests are performed using the reference method on day one as a control.

The in-house repeatability standard deviation is calculated as shown by Formula (5):

$$s_{r,\text{in-house}} = \sqrt{\frac{1}{32}\sum_{j=1}^{8}\sum_{i=1}^{5}\left(y_{ij} - \bar{y}_j\right)^2} \tag{5}$$

where

y_{ij} is the $\log_{10}$ measurement of the i^{th} replicate on day j;

$\bar{y}_j = \frac{1}{5}\sum_{i=1}^{5} y_{ij}$ is the average over the five $\log_{10}$ measurements on day j.

The in-house reproducibility standard deviation is calculated as shown by Formula (6):

$$s_{I,\text{in-house}} = \sqrt{s_A^2 + s_{\text{r,in-house}}^2} \tag{6}$$

where

$$s_A^2 = \frac{1}{7}\sum_{j=1}^{8}\left(\bar{y}_j - \bar{y}\right)^2 - \frac{1}{5}s_{r,\text{in-house}}^2$$

and

$$\bar{y} = \frac{1}{40}\sum_{i=1}^{5}\sum_{j=1}^{8} y_{ij}$$

is the average over all 40 $\log_{10}$ measurements. If $s_A^2 < 0$, set $s_A^2 = 0$.

If the artificially contaminated (food) item is not sufficiently stable, microbiologically, the following options apply.

— Conduct the precision study based on the design described above. The calculated in-house reproducibility standard deviation will very likely be larger than the actual unknown in-house reproducibility standard deviation. Therefore, it can only be used as an upper limit of the actual variation. However, the calculated in-house repeatability standard deviation is not (much) affected by sample instability and can be used without limitation.

— Conduct the precision study based on the design described above, but adjust the results for the underlying trend, see F.2.

— Conduct the precision study based on the design described above but replace the eight days by eight different times. These eight different times shall comprise at least three different days, and within each day, different equipment or different batches/suppliers of media shall be used.

— Conduct the precision study and adjust the results of the alternative method by the results of the reference method, see F.3.

The design of the precision study shall be described in the validation study report.

NOTE 1 The variance components can be calculated more efficiently, with better precision, using the restricted maximum likelihood (REML) approach[13].

NOTE 2 Evaluation of the calculated in-house reproducibility standard deviation can be conducted using the criterion proposed by Reference [11]: target value ±0,25 $\log_{10}$ cfu/g. Since accuracy comprises precision, no additional evaluation of precision is required in this document.

6.2.1.6 Inclusivity/exclusivity study

Inclusivity/exclusivity testing shall be conducted in accordance with 5.2.1.8.

6.2.2 Single-laboratory method validation study without a reference method

6.2.2.1 General

The conventional approach for the single-laboratory validation is based on the ISO 16140-2:2016, 6.1, method comparison study design. In the case of a validation without a reference method, the study has to rely on known contamination levels of artificially contaminated samples.

6.2.2.2 Relative trueness study

The relative trueness study is determined using the known contamination levels of the samples. It shall be conducted in accordance with ISO 16140-2:2016, 6.1.2, where each test result of the reference method is replaced with the known contamination level of the artificially contaminated sample.

6.2.2.3 Accuracy profile study

The accuracy profile study shall be conducted in accordance with ISO 16140-2:2016, 6.1.3, where each test result of the reference method is replaced with the known contamination level of the artificially contaminated sample. In addition, the following adaptations shall be made.

a) The calculation of the standard deviation of the reference method in Step 5 of ISO 16140-2:2016, 6.1.3.3, can be omitted, and the value replaced by 0 since the contamination levels are known and, hence, there is no variability.

b) Step 9 of ISO 16140-2:2016, 6.1.3.3, does not apply because the standard deviation of the reference values cannot exceed the threshold of 0,125.

6.2.2.4 Limit of quantification study

The LOQ study, if applicable, shall be conducted in accordance with ISO 16140-2:2016, 6.1.4.

6.2.2.5 In-house precision study

One (food) item is selected per (food) category for the determination of the in-house repeatability standard deviation and the in-house reproducibility standard deviation.

If the artificially contaminated (food) item is sufficiently stable, microbiologically, tests are conducted on eight different days by three or more technicians. No technician shall conduct tests on more than three days. However, if only two technicians are available, then each technician shall conduct tests on four days. If only one technician is available, this technician shall conduct all tests, but the method is then validated only for this technician.

On each day, the technician performs tests on five replicates, giving a total of 8 × 5 = 40 tests over the eight days.

The in-house repeatability standard deviation is calculated as shown by Formula (7):

$$s_{r,\text{in-house}} = \sqrt{\frac{1}{32}\sum_{j=1}^{8}\sum_{i=1}^{5}\left(y_{ij} - \bar{y}_j\right)^2} \tag{7}$$

where

y_{ij} is the $\log_{10}$ value of the test result of the i^{th} replicate on day j;

$\bar{y}_j = \frac{1}{5}\sum_{i=1}^{5} y_{ij}$ is the average over the five test results on day j.

The in-house reproducibility standard deviation is calculated as shown by Formula (8):

$$s_{I,\text{in-house}} = \sqrt{s_A^2 + s_{r,\text{in-house}}^2} \tag{8}$$

where

$$s_A^2 = \frac{1}{7}\sum_{j=1}^{8}\left(\bar{y}_j - \bar{y}\right)^2 - \frac{1}{5}s_{r,\text{in-house}}^2$$

and

$$\bar{y} = \frac{1}{40}\sum_{i=1}^{5}\sum_{j=1}^{8} y_{ij}$$

is the average over all 40 $\log_{10}$ measurements. If $s_A^2 < 0$, set $s_A^2 = 0$.

If the artificially contaminated (food) item is not sufficiently stable, microbiologically, the following options apply.

— Conduct the precision study based on the design described above. The calculated in-house reproducibility standard deviation will very likely be larger than the actual unknown in-house reproducibility standard deviation. Therefore, it can only be used as an upper limit of the actual variation. However, the calculated in-house repeatability standard deviation is not (much) affected by sample instability and can be used without limitation.

— Conduct the precision study based on the design described above but adjust the results for the underlying trend, see F.2.

— Conduct the precision study based on the design described above, but replace the eight days by eight different times. These eight different times shall comprise at least three different days. Different equipment or different batches/suppliers of media shall be used for each day.

The design of the precision study shall be described in the validation study report.

In-house repeatability and in-house reproducibility standard deviations are used to describe precision of the alternative method. If the in-house reproducibility standard deviation is larger than 0,5, the method cannot be considered to be validated.

NOTE 1 The variance components can be calculated more efficiently, with better precision, using the REML approach.

NOTE 2 Evaluation of the calculated in-house reproducibility standard deviation can be conducted using the criterion proposed by Reference [11]: target value ±0,25 $\log_{10}$ cfu/g. Since accuracy comprises precision, no additional evaluation of precision is required in this document.

6.2.2.6 Inclusivity/exclusivity study

Inclusivity/exclusivity testing shall be conducted in accordance with 5.2.1.8.

7 Summary of acceptability limits

Table 8 provides the overview of acceptability limits (ALs). Criteria applied in the inclusivity/exclusivity study are not included. The organization of the table is based on the following method and study characteristics:

— qualitative or quantitative method;

— factorial or conventional study design;

— with or without reference method;

— paired or unpaired study.

Table 8 — Overview of acceptability limits

Study	Method	Reference method	Paired/ unpaired	Acceptability limit (AL)
Factorial	Qualitative	Yes	Unpaired	(ND–PD) ≤ AL in accordance with Table 5
				RLOD ≤ 2,5
				Factorial difference of RLOD: |d| < 0,6
			Paired	(ND–PD) and (ND+PD) ≤ AL in accordance with Table 5
				RLOD ≤ 1,5
				Factorial difference of RLOD: |d| < 0,6
		No	Not applicable	ND ≤ 3
				PD ≤ 1
				Factorial difference of LOD: |d| < 0,6
	Quantitative	Yes	Paired/ unpaired	Accuracy profile in accordance with ISO 16140-2:2016, 6.1.3.3
		No	Not applicable	Accuracy profile in accordance with ISO 16140-2:2016, 6.1.3.3
Conventional	Qualitative	Yes	Unpaired	(ND–PD) ≤ AL in accordance with ISO 16140-2:2016, 5.1.3.4
				RLOD ≤ 2,5
			Paired	(ND–PD) and (ND+PD) ≤ AL in accordance with ISO 16140-2:2016, 5.1.3.4
				RLOD ≤ 1,5
		No	Not applicable	ND ≤ 3
				PD ≤ 1
	Quantitative	Yes	Paired/ unpaired	Accuracy profile in accordance with ISO 16140-2:2016, 6.1.3.3
		No	Not applicable	Accuracy profile in accordance with ISO 16140-2:2016, 6.1.3.3

Annex A
(informative)

List of factors and factor levels for factorial method validation

The following list can be used to select factors and factor levels for the factorial study designs provided in this document. Factors and factor levels that are expected to have a noticeable effect on the end result should be selected from the following list.

- Culture media and incubation:
  - supplier of non-selective/primary broth;
  - in-house versus ready-made non-selective/primary broth;
  - supplier of selective/secondary broth;
  - in-house versus ready-made selective/secondary broth;
  - supplier of agar plate;
  - in-house versus ready-made agar plate;
  - storage time of broths or plates after incubation before subsequent subculturing or counting (as long as specified within the procedure of the method);
  - age of the medium (within the expiry period of the medium);
  - batches of media (e.g. media for *Enterobacteriaceae* containing bile salts);
  - time of incubation of different steps (e.g. 20 h versus 28 h when 24 h ± 4 h is allowed).
- Background microbiota [at least two levels of background microbiota per (food) item, e.g. fresh and at the expiry of the best-before date]:
  - background microbiota associated with storage (refrigerated/room temperature, stored/not stored);
  - contamination levels (add different levels artificially);
  - interference organisms;
  - batches of (food) items, where background microbiota is known to vary.
- Sample preparation and storage:
  - blender or vortex;
  - resuscitation procedure [e.g. thawing of frozen samples (“a”) at room temperature or (“b”) at 4 °C];
  - immediate use or freezing/thawing of samples;
  - transport conditions: storage or transport of samples.
- Equipment and technicians:
  - technician (use technicians with different level of experience, if possible);
  - different pipetting systems;

  - different incubation conditions [use two different types of incubators, e.g. different 37 °C air incubators, air incubator versus water bath incubation; if only one type of incubator is available, vary incubation conditions at two levels, e.g. level "a" = shortest time permitted by the method tolerance and lowest position in the incubator (= lowest temperature), and level "b" = longest time permitted by the method tolerance and highest position in the incubator (= highest temperature)];
  - spread plates versus pour plates, if relevant.
- Stress factors (if applicable):
  - low/normal temperature;
  - high/normal temperature;
  - chemical stress (e.g. smoking and curing);
  - pH-stress;
  - high pressure stress;
  - low-dose radiation/no radiation (e-beam);
  - dry/non-dry, e.g. low a_w (≤0,60)/high a_w (≥0,95).

NOTE A setting is a specific combination of levels for all factors. Two examples of a setting for the factors from Table A.1 are:

- setting 1 (a,a,b,b): technician "a"; in-house agar plates; refrigerated storage; incubation condition "b";
- setting 2 (b,b,a,a): technician "b"; ready-made agar plates; no storage; incubation condition "a".

Table A.1 — Examples of factors and their levels

Factor	Level "a"	Level "b"
Technician	Technician "a"	Technician "b"
Preparation of agar plates	Made in-house	Purchased ready-made
Storage of inoculated samples (increases background levels)	None	Hold refrigerated 2 days
Incubation condition	Incubation condition "a"	Incubation condition "b"

Annex B
(informative)

Calculation of in-house reproducibility for qualitative methods based on the LOD_{50} study described in 6.1.2.3

The in-house reproducibility parameter captures important information about the performance of methods in microbiology and should be determined in the course of validation studies. The risk with not considering the in-house reproducibility is that the actual reliability of a method is not known and that reliance is placed on inaccurate routine laboratory determinations. This can have serious practical consequences.

For instance, if the validation study takes place on a "good" day, in the sense that the LOD_{50} of the alternative method is found to be 3 cfu/test portion, the true unreliability of the method may go undetected. When a negative result is obtained during routine laboratory determination, it would be concluded that the contamination level is below 3 cfu/test portion. But on a "bad" day, the actual level could have been 30 cfu/test portion. If, however, the validation study had taken place on a "bad" day and the LOD_{50} of the alternative method was found to be 30 cfu/ml, the method would not be validated, even though both methods may be considered, on average, to be equivalent.

The in-house reproducibility standard deviation across (food) items can be estimated as a measure of the precision of the alternative method under different conditions and across (food) items in the same laboratory by the following procedure.

a) Calculate, for each (food) item/day i = 1, 2,..12, the LOD_{50} for the alternative method using a CLL model for the POD. Denote the $\log_{10}$ value of the LOD_{50} for day i by Y_i.

b) Calculate the standard deviation of the $\log_{10}$ value of the LOD_{50}, as shown by Formula (B.1):

$$s_Y = \sqrt{\frac{1}{11}\sum_{i=1}^{12}(Y_i - \bar{Y})^2} \quad \text{(B.1)}$$

where $\bar{Y} = \frac{1}{12}\sum_{i=1}^{12} Y_i$ denotes the average of the $\log_{10}$ value of the LOD_{50}.

c) Extract the standard error of the LOD_{50} for (food) item/day i, (s_i). This value is calculated automatically by standard software for fitting the CLL model.

d) Calculate the average standard error, as shown by Formula (B.2):

$$s_0 = \sqrt{\frac{1}{12}\sum_{i=1}^{12} s_i^2} \quad \text{(B.2)}$$

e) Calculate the in-house reproducibility standard deviation across (food) items, as shown by Formula (B.3):

$$s_{\text{in-house}} = \sqrt{s_Y^2 - s_0^2} \quad \text{(B.3)}$$

The in-house precision is set at 0 when the difference in step e) is negative.

An in-house reproducibility standard deviation, $s_{\text{in-house}}$, greater than 0,5 indicates that there is considerable variation of LOD_{50} between different (food) items or days.

Annex C
(informative)

Example of a factorial single-laboratory method validation study for a quantitative method against a reference method

C.1 General

This annex provides an example to illustrate the procedure for the performance of the validation of a quantitative method against a reference method as described in 5.2.

C.2 Study design

Twelve food items from three food types of the food category "raw milk and dairy products" were selected. Samples were artificially contaminated at three contamination levels: a low level L_1, a medium level L_2 and a high level L_3 in accordance with Table C.1.

Table C.1 — Food items and contamination levels

Level	Food item
Low	1, 4, 7, 10
Medium	2, 5, 8, 11
High	3, 6, 9, 12

Two technicians performed the reference method (see ISO 4833-1) and the alternative method for the aerobic plate count of the dairy products. Incubation times, for the alternative method and the reference method, were 48 h ± 3 h and 72 h ± 3 h, respectively.

Each food item was analysed in two settings, and in each setting two replicates were conducted. Therefore, 48 (12 food items × 2 settings × 2 replicates) single tests were conducted per method. A total of 96 (48 reference and 48 alternative) single tests were conducted, with each technician performing 48 tests.

Table C.2 shows the allocation of samples to the settings of factor level combinations.

Four influencing factors were evaluated in the study:

— technician;

— culture medium;

— incubation condition;

— incubation time.

All twelve food items were evaluated in a paired study design: the same replicates were used for both the alternative and the reference method. The variations in technicians, culture medium, incubation conditions and incubation time in accordance with the experimental design summarized in Table C.2 were applied to both the alternative and the reference method.

Detailed instructions were provided for the analysis, including a flowchart for each of the two technicians (see Figures C.1 and C.2).

Table C.2 — Experimental design

Setting	Food item	Factor 1 Technician	Factor 2 Culture medium	Factor 3 Incubation condition	Factor 4 Incubation time	
					Alternative	Reference
1	1, 2, 3	a	Pre-made dilution buffer	Incubation condition A	45 h	69 h
2	1, 2, 3	b	Dehydrated dilution buffer	Incubation condition B	51 h	75 h
3	4, 5, 6	a	Dehydrated dilution buffer	Incubation condition A	51 h	75 h
4	4, 5, 6	b	Pre-made dilution buffer	Incubation condition B	45 h	69 h
5	7, 8, 9	a	Dehydrated dilution buffer	Incubation condition B	45 h	69 h
6	7, 8, 9	b	Pre-made dilution buffer	Incubation condition A	51 h	75 h
7	10, 11, 12	a	Pre-made dilution buffer	Incubation condition B	51 h	75 h
8	10, 11, 12	b	Dehydrated dilution buffer	Incubation condition A	45 h	69 h

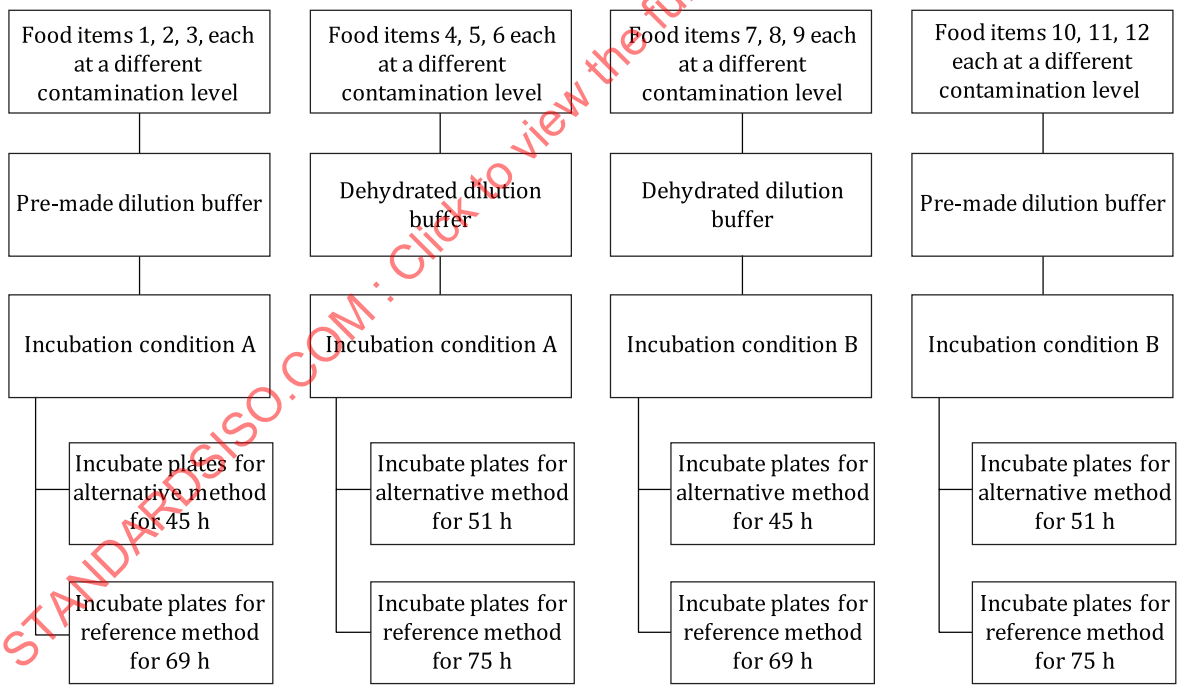


Figure C.1 — Instructions for technician "a"

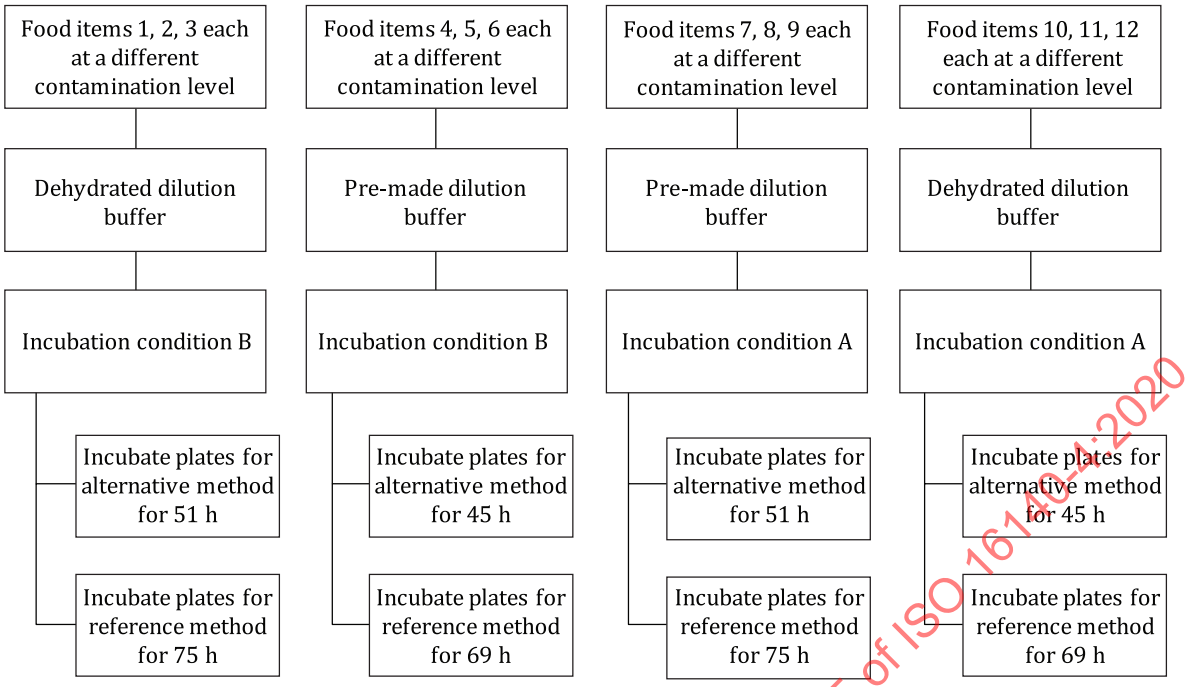


Figure C.2 — Instructions for technician "b"

C.3 Calculations and summary of data

C.3.1 Summary of the results

Upon receipt of the data, it was verified that both technicians performed the analyses in accordance with the methods as indicated in the standard operation procedure. No outlier tests were performed on the data, and no data points were excluded.

All single test result calculations were $\log_{10}$ transformed. The results are given in Table C.3.

C.3.2 Relative trueness

The calculation of relative trueness was conducted in accordance with the steps described in 5.2.1.5. A summary of the results is given in Table C.4.

C.3.3 Accuracy profile

The calculation of the accuracy profile was conducted in accordance with the steps described in 5.2.1.6, based on the individual test results tabulated in Table C.5. Figure C.3 is the graphical representation of the accuracy profile.

Table C.3 — Results of single-laboratory method validation study

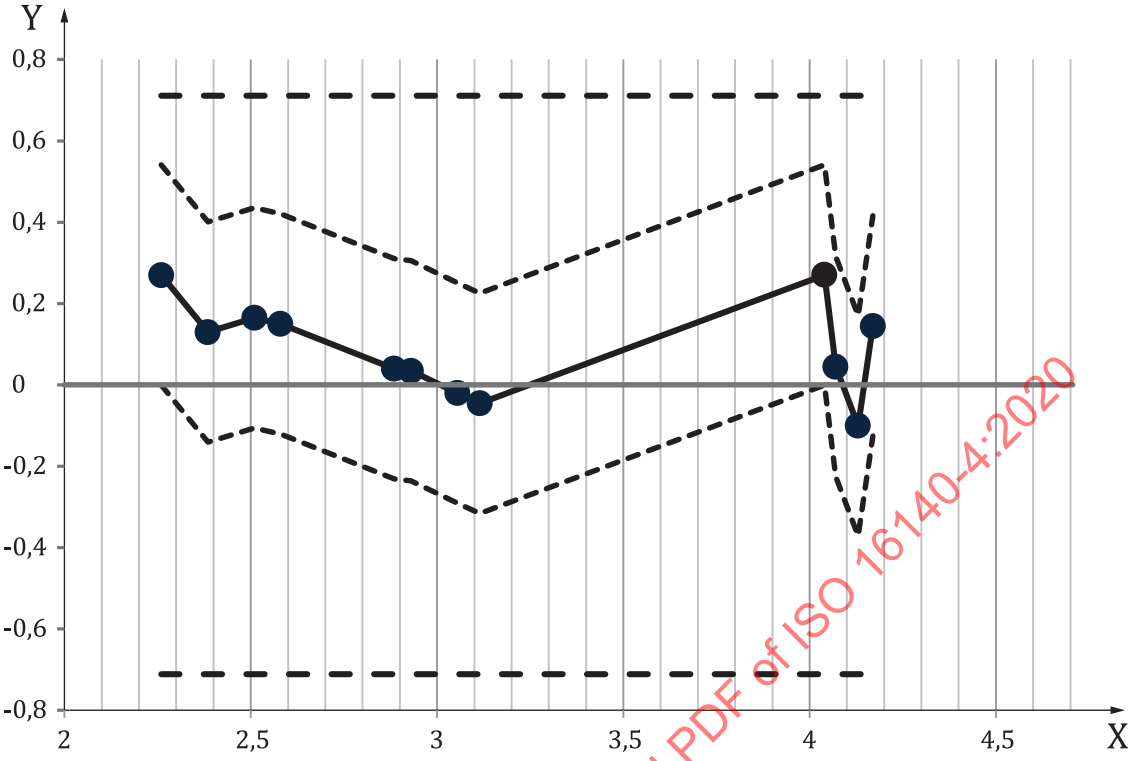
Contamination level	Setting	Food item	log_{10} cfu/g or ml – Single test results			
			Reference		Alternative	
			Replicate 1	Replicate 2	Replicate 1	Replicate 2
Low	1	1	2,08	2,11	2,65	2,93
	2		2,52	2,41	2,41	2,38
	3	4	2,69	2,36	2,73	2,84
	4		2,62	2,40	2,62	2,40
	5	7	2,15	2,45	2,76	2,54
	6		2,45	2,32	2,30	2,49
	7	10	2,53	2,62	2,82	2,74
	8		2,54	2,78	2,51	2,72
Medium	1	2	3,28	3,26	3,20	3,20
	2		2,97	2,91	2,94	2,91
	3	5	2,83	2,93	2,91	3,11
	4		2,93	2,98	2,96	2,97
	5	8	3,36	3,18	3,08	3,15
	6		2,91	2,93	2,99	2,90
	7	11	2,89	2,88	2,81	2,66
	8		2,94	2,86	3,04	3,08
High	1	3	3,94	3,79	4,11	4,12
	2		4,34	4,20	4,15	4,08
	3	6	4,26	4,23	4,56	4,51
	4		4,11	3,89	4,12	4,11
	5	9	4,15	3,72	3,98	3,74
	6		4,34	4,11	4,30	4,08
	7	12	4,08	3,80	4,62	4,36
	8		4,18	4,00	4,26	4,08

Table C.4 — Summary of relative trueness

Contamination level	Food item	$\log_{10}$ cfu/g or ml			
		Mean test results		Mean	Difference
		Reference	Alternative		
Low	1	2,280	2,593	2,436	0,313
	4	2,518	2,648	2,583	0,130
	7	2,343	2,523	2,433	0,180
	10	2,618	2,698	2,658	0,080
Medium	2	3,105	3,063	3,084	−0,043
	5	2,918	2,988	2,953	0,070
	8	3,095	3,030	3,063	−0,065
	11	2,893	2,898	2,895	0,005
High	3	4,068	4,115	4,091	0,048
	6	4,123	4,325	4,224	0,203
	9	4,080	4,025	4,053	−0,055
	12	4,015	4,330	4,173	0,315
Average differences: Total					0,098
Standard deviation of differences: Total					0,133
Average differences: Technician “a”					0,190
Average differences: Technician “b”					0,007
Average differences: Pre-made dilution buffer					0,134
Average differences: Dehydrated dilution buffer					0,063
Average differences: Incubation condition A					0,146
Average differences: Incubation condition B					0,051
Average differences: Incubation time short					0,110
Average differences: Incubation time long					0,087

Table C.5 — Tabulated results for the accuracy profile

Contamination level	Food item	Setting	$\log_{10}$ cfu/g or ml – Single test results							
			Reference				Alternative			
Low	1	1, 2	2,08	2,11	2,52	2,41	2,65	2,93	2,41	2,38
	4	3, 4	2,69	2,36	2,62	2,40	2,73	2,84	2,62	2,40
	7	5, 6	2,15	2,45	2,45	2,32	2,76	2,54	2,30	2,49
	10	7, 8	2,53	2,62	2,54	2,78	2,82	2,74	2,51	2,72
Medium	2	1, 2	3,28	3,26	2,97	2,91	3,20	3,20	2,94	2,91
	5	3, 4	2,83	2,93	2,93	2,98	2,91	3,11	2,96	2,97
	8	5, 6	3,36	3,18	2,91	2,93	3,08	3,15	2,99	2,90
	11	7, 8	2,89	2,88	2,94	2,86	2,81	2,66	3,04	3,08
High	3	1, 2	3,94	3,79	4,34	4,20	4,11	4,12	4,15	4,08
	6	3, 4	4,26	4,23	4,11	3,89	4,56	4,51	4,12	4,11
	9	5, 6	4,15	3,72	4,34	4,11	3,98	3,74	4,30	4,08
	12	7, 8	4,08	3,80	4,18	4,00	4,62	4,36	4,26	4,08



Key

X levels $\log_{10}$ cfu
Y accuracy (difference of log)
absolute bias
upper and lower β-ETI
line of zero bias
upper and lower acceptability limits

Figure C.3 — Accuracy profile

C.3.4 Precision data

The calculation of precision data was conducted in accordance with the steps described in 5.2.1.7. Table C.6 shows the tabulated results for the precision data. Precision for the alternative method was quite close to the precision of the reference method.

Table C.6 — Tabulated results for the precision data

	Reference	Alternative
$s_{r,\text{in-house}}$	0,132	0,109
$s_{L,\text{in-house}}$	0,146	0,181
$s_{R,\text{in-house}}$	0,197	0,211

C.3.5 Interpretation

Acceptability criterion of the accuracy profile is fulfilled. No critical factorial effects were detected, and precision of the alternative method is similar to the precision of the reference method.

Annex D
(informative)

Example of a factorial single-laboratory method validation study for a qualitative method against a reference method

This annex provides an example to illustrate the procedure for the computation of validation parameters in accordance with 5.1. The study is unpaired. The experimental design provided in Table 3 is implemented. The tests were carried out with twelve (food) items allocated to four blocks, each block consisting of two settings, at two contamination levels.

In this example, the food category "heat-processed milk and dairy products" is validated by choosing, for example, the three food types "pasteurized dairy products", "sterilized or UHT dairy products" and "dry".

The study design is presented in Table D.1. The test results for both the alternative and the reference methods are provided in Table D.2. Note that in this example, the results of the alternative method and the confirmed alternative method are identical. Therefore, for the sake of simplicity, no distinction is made between the results of the alternative method and the results of the confirmed alternative method. In addition, six zero level L_0 (food) items tested negative with both the reference and the alternative methods.

Table D.1 — Design for both the alternative and the reference methods (fractional and high contamination level)

Food items 1 to 12		**Block 1: 1 to 3**		**Block 2: 4 to 6**		**Block 3: 7 to 9**		**Block 4: 10 to 12**	
Setting		1	2	3	4	5	6	7	8
Factor 1	technician	a	b	a	b	a	b	a	b
Factor 2	culture medium	b	a	a	b	a	b	b	a
Factor 3	background microbiota	a	b	a	b	b	a	b	a
Factor 4	incubation condition	a	b	b	a	a	b	b	a