INTERNATIONAL STANDARD

ISO 8196-3
IDF 128-3

Second edition
2022-05

Milk — Definition and evaluation of the overall accuracy of alternative methods of milk analysis —

Part 3:
Protocol for the evaluation and validation of alternative quantitative methods of milk analysis

Lait — Définition et évaluation de la précision globale des méthodes alternatives d'analyse du lait —

Partie 3: Protocole d'évaluation et de validation des méthodes quantitatives alternatives pour l'analyse du lait

Reference numbers
ISO 8196-3:2022(E)
IDF 128-3:2022(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

International Dairy Federation
Silver Building • Bd Auguste Reyers 70/B
B-1030 Brussels
Phone: +32 2 325 67 40
Fax: +32 2 325 67 41
Email: info@fil-idf.org
Website: www.fil-idf.org

Published in Switzerland

Contents

Page

Forewords

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 5, *Milk and milk products*, and the International Dairy Federation (IDF). It is being published jointly by ISO and IDF.

This second edition cancels and replaces the first edition (ISO 8196 | IDF 128-3:2009), which has been technically revised. The main changes are as follows:

— the validation scheme has been simplified for phase II and it is possible to validate a new instrument with the comparison with a previous validated instrument.

A list of all parts in the ISO 8196 | IDF 128 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.

IDF (the International Dairy Federation) is a non-profit private sector organization representing the interests of various stakeholders in dairying at the global level. IDF members are organized in National Committees, which are national associations composed of representatives of dairy-related national interest groups including dairy farmers, dairy processing industry, dairy suppliers, academics and governments/food control authorities.

ISO and IDF collaborate closely on all matters of standardization relating to methods of analysis and sampling for milk and milk products. Since 2001, ISO and IDF jointly publish their International Standards using the logos and reference numbers of both organizations.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. IDF 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.

This document was prepared by the IDF *Standing Committee on Statistics and Automation* and ISO Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 5, *Milk and milk products.* It is being published jointly by ISO and IDF.

The work was carried out by the IDF/ISO Action Team (S14) of the *Standing Committee on Statistics and Automation* under the aegis of its project leader, Dr S. Orlandini (IT).

Introduction

This document is complementary to ISO 8196-1 | IDF 128-1. It describes a protocol for the evaluation of new alternative methods for which ISO 8196-1 | IDF 128-1 cannot apply, e.g. when the organization of interlaboratory studies is hampered by a limited number of new instruments available for study.

The latter is generally the case with dedicated instrumental methods (e.g. milk payment analysis, milk recording analysis) of which the commercialization depends on official approvals for use. An application for such an official approval is to be accompanied by one or more assessments of the relevant performance characteristics.

This document specifies a harmonized protocol for such a method validation by expert laboratories. It lists the evaluation steps and provides a criteria-based approach for the assessment of the performance characteristics, including guidance for checking statistical compliance.

On the basis of such a harmonized protocol, a limited number of evaluations should suffice for a decision by an approval body for the application of the method and/or equipment. Examples with indicative limits are given for the evaluation of a method for the determination of fat, protein, lactose, urea and somatic cell count in milk. The guideline can also be applied to other parameters such as freezing point and pH in milk.

Milk — Definition and evaluation of the overall accuracy of alternative methods of milk analysis —

Part 3:
Protocol for the evaluation and validation of alternative quantitative methods of milk analysis

1 Scope

This document specifies a protocol for the evaluation and validation of alternative quantitative methods of milk analysis. This document is also applicable for the validation of new alternative methods where, due to a limited number of operational instruments, the execution of an interlaboratory study and ISO 8196-1 | IDF 128-1 is not feasible.

The protocol is applicable to milk parameters such as, for example, fat, protein, lactose, urea and somatic cells in milk. It can also be extended to other parameters.

This document also establishes the general principles of a procedure for granting international approvals for the performance of the alternative methods. These principles are based on the validation protocol defined in this document.

2 Normative references

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

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

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

ISO 8196-1 | IDF 128-1, *Milk — Definition and evaluation of the overall accuracy of alternative methods of milk analysis — Part 1: Analytical attributes of alternative methods*

ISO 8196-2 | IDF 128-2, *Milk — Definition and evaluation of the overall accuracy of alternative methods of milk analysis — Part 2: Calibration and quality control in the dairy laboratory*

ISO/IEC 17025, *General requirements for the competence of testing and calibration laboratories*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534-1, ISO 5725-1, ISO 8196-1 | IDF 128-1, ISO 8196-2 | IDF 128-2 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 https://www.electropedia.org/

3.1
validation of alternative method
verification of the performance of an alternative method on whether it is adequate for the intended use

3.2
measurand
quantity intended to be measured

Note 1 to entry: A measurand may be a milk component (e.g. fat and protein), a physical characteristic (e.g. freezing point) or a biological element (e.g. somatic cells).

Note 2 to entry: Adapted from ISO/IEC Guide 99:2007, 2.3.

3.3
quantitative method
method of analysis whereby the result is an amount of a quantity, a concentration or a value of a *measurand* (3.2) determined either directly or on a test portion

3.4
methods comparison study
study performed by an *expert laboratory* (3.6) of an alternative method against the reference method or a comparison method/instrument under test bed conditions

3.5
interlaboratory study
study of the performance of an alternative method on one or more "identical" laboratory samples of homogeneous, stable materials under documented conditions in several laboratories and under the control of an *organizing laboratory* (3.7)

Note 1 to entry: The data interpretation should be performed in collaboration with *expert laboratory* (3.6).

3.6
expert laboratory
laboratory having qualified staff and equipment to perform a *methods comparison study* (3.4)

Note 1 to entry: The expert laboratory is specialized in analytical evaluations and shall conform to ISO/IEC 17025 as well as having relevant experience in the area of application.

3.7
organizing laboratory
laboratory having staff with statistical expertise and qualified staff and necessary equipment to prepare the samples to perform an *interlaboratory study* (3.5)

Note 1 to entry: The organizing laboratory shall operate in conformity with ISO/IEC 17025 for the method used to check the homogeneity of the samples.

3.8
national approval
authorization of the use of a method for defined purposes in a country, generally for reasons of collective interest and/or having an official character, delivered by an approval body

3.9
international approval
authorization of the use of a method for defined purposes at international level, generally for reasons of collective interest and/or having an official character, delivered by an approval body for the benefit of stakeholders

4 General principles for the validation of alternative methods

4.1 Validation protocol

4.1.1 General

The validation protocol comprises two phases as specified in 4.1.2 and 4.1.3.

4.1.2 Phase I

A methods comparison study includes the assessment of the performance characteristics of the alternative method under validation. A comparison of the alternative method against the reference method under test bed conditions is required. In cases where the instrument under evaluation has the same analytical principle and only minor technical changes from the previously validated version, the comparison can be done between the two instruments, considering the results of the oldest version as an anchor to evaluate the results of the new instrument generation.

This part of the evaluation shall be carried out by an expert laboratory.

4.1.3 Phase II

A method confirmation study under routine testing conditions is initiated after a successful Phase I. It is recommended to examine at least two instruments, for national approval, or three instruments, for international approval.

Depending on the purpose, the approval body can decide whether two or three instruments are to be examined and whether the instruments are to be located in the same laboratory or in different laboratories and geographies under routine testing conditions. A test period of a minimum of two months is recommended for Phase II or to organize an interlaboratory study associated with the data collection from routine laboratories. For this phase, detailed steps are described in 5.3.2.

4.1.4 National approval

Based on the content of submitted reports, a competent body can grant a national approval, indicating sufficient quality in measurement results and adequateness of the alternative method for the proposed purpose.

4.1.5 International approval

Approval bodies or international organizations can grant an international approval. International approval can be granted based on three single national validations or the results of Phase I performed in an expert laboratory and the results from a method confirmation study or an interlaboratory study as described in 4.1.3.

4.2 Field of validity of the approval

This protocol is applicable to the validation of alternative methods for the quantitative compositional analysis and somatic cell count determination in raw milk from cow, sheep, goat and buffalo. The validation study shall be conducted separately for the milk of each species. When a component under validation occurs with unusual concentrations (e.g. Jersey breed with high fat and protein content) the evaluation should be carried out over the whole relevant range of the concerned component.

The method and/or instrument should be evaluated with the configuration as offered by the concerned manufacturer. If the configuration changes, it should be proven in an independent way that it does not influence the precision and the accuracy beyond acceptable limits.

Carefully note and report all characteristics of both the milk products analysed, the calibration model(s) version and the configuration(s) of the alternative method assessed.

5 Technical protocol for the validation

5.1 Course of operations

Whatever the alternative method, a standard measurement process can be represented schematically as shown in Figure A.1. Each step corresponds to a source of error that can contribute to the overall uncertainty of the method. The evaluation protocol and experimental designs are constructed to fit the sequence of signal treatment and to permit verification that they are set up in such a way that precision and accuracy of the method can respond to the limits required in practice.

It is necessary for each step of the evaluation described in the following paragraphs to fulfil the appropriate limits for each analytical criterion before starting the next step.

The methods comparison study (Phase I) defines the minimum assessment sequence to be carried out.

The method confirmation/interlaboratory study (Phase II) provides complementary information on the method performance under routine use conditions.

5.2 Methods comparison study (Phase I)

5.2.1 General

The evaluation is to be carried out with test results expressed in standardized units of the reference method. For methods covering large ranges of measured values (i.e. wider than one log unit), it is recommended to split the range into levels, each of maximum width one log unit, so as to obtain a minimum of three levels and to perform statistical calculations separately on each level. Where appropriate, a logarithmic transformation of the data can be applied, see 5.2.2.

NOTE 1 For instance, for fat in commercial milk, distinction can be made between skim milk, semi-skimmed milk and whole milk. For raw milk, natural fat and protein ranges are often related to the species, which are then to be assessed by separate evaluations (see 4.2). Somatic cells in raw milk typically cover a range of several log units.

Evaluation results should conform to the specifications stated in the following paragraphs. For general dairy industry purposes, limits for the different analytical characteristics mentioned have been extracted or derived from existing International Standards.

Annex B summarizes these limits for fat, protein (crude protein, true protein and casein), lactose, urea, somatic cells, freezing point and pH as indicative limits obtained from proficiency tests.

NOTE 2 For liquid milk during milking or processing, there can be different assessment criteria for in-line and at-line analyses systems.

5.2.2 Compulsory assessments for the validation

5.2.2.1 Assessment of preliminary instrumental fittings

5.2.2.1.1 General

Before starting any further assessment, basic criteria indicating a proper functioning of the method or the instrument require verification. These criteria are repeatability, intralaboratory reproducibility carry-over and linearity.

5.2.2.1.2 Precision (repeatability and intralaboratory reproducibility)

The method used should present a stable measurement signal that conforms to the precision requirements. If not, the analyser is either not functioning correctly (and should not be used) or its precision is not appropriate for the objective of the analysis. Hence, the instantaneous stability (repeatability) and the signal level stability shall be assessed prior to any other characteristics.

The precision should be evaluated at three different concentration levels of the component measured: low, medium, and high.

During the day, analyse pilot milk samples in triplicate (n = 3) every 15 min to 20 min of instrument activity without any change in the calibration in order to obtain results from a minimum of 20 pilot samples analysed for each level ($q \geq 20$). Preferably, the instrument should be operated under conditions as close as possible to routine circumstances. Sufficient numbers of samples should be processed to keep the instrument running between the periodic checks.

Estimate for each pilot:

a) s_r, the standard deviation of repeatability;

b) s_p, the standard deviation of mean pilots;

c) s_c, the standard deviation between time periods;

d) $s_{R\text{intra}}$, the standard deviation of intralaboratory reproducibility.

For each time period (i = 1,2, ... q), calculate the pilot sample mean $\overline{x_j}$ and the standard deviation of the mean pilot s_j over the q replicate measurements, as shown by Formulae (1) and (2):

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \tag{1}$$

$$s_j = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2\right)^{1/2} \tag{2}$$

where

n is the number of replicates at each time period (typically n = 3).

The overall repeatability standard deviation of this pilot is found by averaging these s_j^2 over all the q time periods in the day, as shown by Formula (3):

$$s_r = \left(\frac{1}{q}\sum_{j=1}^{q} s_j^2\right)^{1/2} \tag{3}$$

where

q is the number of time periods.

and the standard deviation of mean pilots, as shown by Formula (4):

$$s_p = \sqrt{\left[\frac{1}{q-1}\sum_{j=1}^{q}\left(\bar{x}_j - \bar{x}^2\right)\right]} \tag{4}$$

where $\bar{x} = \frac{1}{q}\sum_{i=1}^{q}\overline{x_i}$

The corrected standard deviation between time periods (for this pilot) is given by Formula (5):

$$s_c = (s_b^2 - s_r^2/n)^{1/2} \tag{5}$$

with $s_c = 0$ if $s_c < 0$.

The overall standard deviation of intralaboratory reproducibility for this pilot is shown by Formula (6):

$$s_{R\text{intra}} = \sqrt{s_r^2 + s_c^2} \tag{6}$$

The values obtained for s_R and $s_{R\text{intra}}$ should conform to the limits stated in Annex B.

The stability of the method response during the analyses of the pilot sample can be visualized by plotting the means $\bar{x}_j$ of the different three pilots means versus the time. See the example in Clause C.1.

5.2.2.1.3 Carry-over effect

5.2.2.1.3.1 Strong differences in component concentrations between two successively analysed samples can influence the result of the second.

Differences can be caused by incomplete rinsing of the flow system and the measuring cell by liquid circulation and contamination by the stirring device. Automatic correction of results is acceptable within certain limits, provided it can be proven that there is a systematic transfer of a small quantity of material from one measurement to the next.

Automated analysers for liquids often allow automatic correction to compensate for the overall carry-over effect when necessary. Carry-over shall be clearly distinguished from rinsing efficiency.

5.2.2.1.3.2 The overall carry-over effect should be assessed including the correction factors either set in the instrument or obtained using the method supplied by the manufacturer. It should not exceed the values stated per component.

Limits are defined from the prerequisite that carry-over effect should not produce an error higher than the repeatability of the method. Hence, limits for the carry-over ratio (COR), L_C, should fulfil the condition $L_C \leq (r/\Delta L_{range}) \times 100$ where r is the repeatability limit at the level of the bias measured and ΔL_{range} is the difference between the maximum and the minimum concentration in the range of interest. For components where repeatability is not constant over the measuring range, the COR limits are set based on the levels of best repeatability (e.g. somatic cell counting). Common limits for COR are in the range 1 % to 2 %.

5.2.2.1.3.3 The rinsing efficiency of the flow system shall be assessed separately by running tests without any correction (correction factor set to zero) in manual mode that bypasses the stirrer. The carry-over should not exceed 1 % as given in ISO 9622 | IDF 141 or 2 % as given in ISO 13366-2 | IDF 148-2.

5.2.2.1.3.4 Analyse two samples, with high and low concentrations of prior distribution in series of test portions. Repeat, as many times, as necessary (see below) the analytical sequence in terms of component concentration, low, low, high, high, in order to obtain N_C sets of results, L_{L1}, L_{L2}, L_{H1} and L_{H2}. The minimum number of sequences, N_C, should be 20.

NOTE For components where repeatability is not constant over the measuring range and for levels with high repeatability, more numerous sequences can be required. Alternative numbers of sequences can be calculated by $N_C \geq [r \times 100/(L_C \Delta L_{test})]^2$ where ΔL_{test} is the range between high and low concentration samples (equal to or greater than ΔL_{range}).

5.2.2.1.3.5 Method requirements for samples: Prepare a sufficient number of test portions from each low and high concentration laboratory sample prior to analysis in order to analyse each test portion only once. The low and high concentration laboratory samples should preferably be milks or liquid products with similar viscosity to those routinely analysed.

Ensure that individual component concentrations differ considerably. For milk, this can, for instance, be achieved by using natural separation (creaming for fat), artificial separation (ultrafiltration for protein, microfiltration for somatic cells) or addition (lactose and urea).

For biochemical component determinations, the low and high concentrations of the laboratory samples should, preferably, be extreme values in the measuring range.

Sufficiently large ranges are recommended to easily differentiate carry-over effects from random error. The minimum range needed, $\Delta L_{test} = L_H - L_L$, can be calculated according to $\Delta L_{test} \geq r \times 100/(L_C\sqrt{N_C})$ where r and L_C are the stated limits and N_C is the number of sequences applied (see Annex B).

For milk components or criteria covering large ranges of concentration, e.g. from 10 to 1 000, the ratio of carry-over error may not be constant over the whole range. This should be verified by assessing the carry-over at different concentrations.

In such case, it is recommended to choose a level L_{Hi} at the median of each part, i, previously defined in the whole range. A minimum number of two levels in the medium and high concentration range is needed that can be extended to three for particularly wide ranges.

Indication for somatic cell counting in individual animal milk, the definition of three levels, at about 500 x 10^3 cells/ml, 1 000 x 10^3 cells/ml and 1 500 x 10^3 cells/ml, is recommended.

5.2.2.1.3.6 Calculation: Calculate the mean of the differences, $d_{LLi} = L_{L1i} - L_{L2i}$ and $d_{LHi} = L_{H2i} - L_{H1i}$, $d_{LL,}$ d_{LH} and the mean difference of concentration, $d_{\acute{A}} = \overline{L}_{H2} - \overline{L}_{L2}$

The COR can be obtained by using Formulae (7) and (8):

$$C_{H/L} = \left(\sum L_{L1} - \sum L_{L2}\right) \times 100 / \left(\sum L_{H2} - \sum L_{L2}\right) = \left(\overline{L}_{L1} - \overline{L}_{L2}\right) \times 100 / \left(\overline{L}_{H2} - \overline{L}_{L2}\right) \quad (7)$$

$$C_{L/H} = \left(\sum L_{H2} - \sum L_{H1}\right) \times 100 / \left(\sum L_{H2} - \sum L_{L2}\right) = \left(\overline{L}_{H2} - \overline{L}_{H1}\right) \times 100 / \left(\overline{L}_{H2} - \overline{L}_{L2}\right) \quad (8)$$

The two should not exceed the limit, L_C, in the test condition stated for the component reported in Annex B.

5.2.2.1.4 Linearity

5.2.2.1.4.1 General

According to the classical definition of an indirect method, the instrument signal should result from a characteristic of the component measured and thereby allow the definition of a simple relationship to the component concentration.

Linearity expresses the constancy of the ratio between the increase in the concentration of a milk component and the corresponding increase of the alternative method result. Therefore, linearity of the measurement signal is in most cases essential to maintain a constant sensitivity over the measuring range and to allow easy handling of calibration and fittings. Moreover, it allows in routine (to some extent) measurements beyond the calibration range through linear extrapolation.

The method is specified in 5.2.2.1.4.2 to 5.2.2.1.4.4.

5.2.2.1.4.2 Samples

Linearity can be assessed using sets of 8 to 15 samples with component concentrations evenly distributed over the measuring range.

a) Samples should preferably be milks or liquids of similar physical characteristics (i.e. density, viscosity), e.g. by combining (weighing) a high content sample, L_H, and a low content sample, L_L.

b) Concentrations should vary in regular intervals. Depending on the component, that can for instance be achieved by natural separation (creaming for milk fat), artificial separation (ultrafiltration for protein, microfiltration for somatic cells) and recombination, or by using pure solutions (lactose