INTERNATIONAL STANDARD

ISO 8196-1

First edition
2000-06-15

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

Part 1: Analytical attributes of indirect methods

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

Partie 1: Attributs d'analyse de méthodes indirectes

Reference number
ISO 8196-1:2000(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.ch
Web www.iso.ch

Printed in Switzerland

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this part of ISO 8196 may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

International Standard ISO 8196-1 was prepared by Technical Committee ISO/TC 34, *Agricultural food products*, Subcommittee SC 5, *Milk and milk products*, in collaboration with the International Dairy Federation (IDF) and AOAC International, and will also be published by these organizations.

ISO 8196 consists of the following parts, under the general title *Milk — Definition and evaluation of the overall accuracy of indirect methods of milk analysis*:

— *Part 1: Analytical attributes of indirect methods*

— *Part 2: Calibration and quality control in the dairy laboratory*

Introduction

The main purpose of this part of ISO 8196 is to give definitions of the various performance characteristics which are covered by the general concept of “overall accuracy” of an analytical method, experimental designs and recommended statistical procedures to evaluate these characteristics quantitatively.

The performance characteristics of an analytical method may be defined as a set of quantitative and experimentally determined values for criteria of fundamental importance in assessing the suitability of a method for any given purpose. The general concepts apply to all analytical methods, but special emphasis is given to rapid indirect physico-chemical methods which are currently in use for testing milkfat, protein, lactose and total solids content.

As an application of accuracy data, ISO 8196-2 gives practical details and recommendations for the calibration of instruments and quality control in routine dairy laboratories.

While this part of ISO 8196 is mainly intended for experts to assess new indirect instrumental methods of analysis, part 2 gives guidance for routine laboratories using these methods.

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

Part 1:
Analytical attributes of indirect methods

1 Scope

This part of ISO 8196 defines the various performance characteristics that constitute the overall accuracy of an analytical method and describes the design of experiments and the recommended procedures to be used to evaluate these characteristics quantitatively. It also gives recommendations for the calibration of instruments and quality control procedures for dairy laboratories.

2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this part of ISO 8196. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this part of ISO 8196 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistics terms.*

ISO 3534-2, *Statistics — Vocabulary and symbols — Part 2: Statistical quality control.*

ISO 3534-3, *Statistics — Vocabulary and symbols — Part 3: Design of experiments.*

ISO 5725 (all parts), *Accuracy (trueness and precision) of measurement methods and results.*

3 Terms and definitions

For the purposes of this part of ISO 8196, the terms and definitions given in parts 1 to 3 of ISO 3534 and ISO 5725-1 apply, together with the following.

3.1 General definitions

3.1.1
true value
that value, known or assigned, which characterizes a measured quantity perfectly defined under the conditions which exist at the moment when that quantity is observed

NOTE It is an ideal value which could be achieved only if all causes of measurement error were eliminated and the population were infinite.

3.1.2
reference method
method internationally recognized by experts or by agreement between the parties, which gives the “true” or “assigned value” of the quantity of the determinant to be measured

3.1.3
indirect method
method which does not measure directly the component that is intended to be measured, but instead measures one or more quantities or properties which are functionally linked to that component

NOTE The signal is related to known quantitative values, or true values, of the component, obtained either by standard materials or instruments, or more often by using a reference method.

3.1.4
overall accuracy
degree of confidence for a measured value, or the extent of its correctness

NOTE 1 This is an index which indicates the amount of error involved, and is usually expressed as the error associated with the method used and calculated under appropriate conditions.

NOTE 2 When a single quantitative measurement (x_i) of a specific determinant (or variable) is made with a given method of analysis, that measurement is always an estimate of its true value (μ). The error of the method is given by the difference $x_i - \mu$. The overall accuracy is best when the difference $x_i - \mu$ is the smallest.

Basically, this difference depends on the following four major analytical characteristics of the method:

— precision,
— accuracy of the mean,
— sensitivity,
— limit of detection.

NOTE 3 Only the precision and accuracy of the mean will be considered in ISO 8196.

3.1.5
sensitivity
smallest change in concentration which can be measured by an analytical procedure

3.1.6
limit of detection
smallest concentration (or amount) of a substance which can be detected with a specified degree of confidence by an analytical procedure

NOTE For indirect instrumental methods, sensitivity and limit of detection are usually determined by the sensitivity of the detector and the signal/noise ratio. These characteristics are generally very good for instrumental methods used for chemical milk analysis.

3.2 Statistical definitions

3.2.1
precision
closeness of agreement between results obtained by a given standard method on identical material under prescribed conditions

3.2.2
repeatability limit
r
value below which the absolute difference between two single test results obtained with the same method on identical test material, under the same conditions (same operator, same apparatus, same laboratory, and in rapid succession) may be expected to lie with a specified probability

NOTE In the absence of any other indication, the probability is 95 %.

3.2.3
reproducibility limit
R
value below which the absolute difference between two single test results on identical material obtained under different conditions (different operators, different apparatus, different laboratories and/or different times) may be expected to lie with a specified probability

NOTE In the absence of any other indication, the probability is 95 %.

3.2.4
accuracy of the mean
closeness of agreement between the true value and the mean of results which would be obtained by applying the test procedure a large number of times in order to minimize precision random errors

NOTE This refers to that portion of the overall error associated with a method at a particular level of the determinant that is not due to the random error of measurement but is attributable to known or unknown factors which prevent the measurement from reaching the true value.

3.2.5
exactness of calibration
closeness of agreement, at each level of the determinant, between the indirect method value and the estimated mean of the true value given by the reference method for all individual samples at the corresponding level

3.2.6
accuracy
closeness of agreement between the individual means of test results obtained by the reference method and the indirect method on identical materials, providing the calibration of the indirect method is exact

4 Statistical expressions

4.1 Precision

Basically precision covers all types of fortuitous and random errors which cannot be completely avoided and whose main characteristics vary from one test to another (e.g. volume delivered by a pipette, environmental conditions, stability of an instrument, electronic noise, etc.).

Mistakes, such as misreadings or operational mistakes or more generally any value found as outlier with the appropriate tests and which are worth considering, are not included in precision data.

Obviously, the variability between test results will be smallest when tests are performed within a laboratory under the most identical conditions (replication conditions), and largest when tests are performed by different laboratories under quite different conditions (repetition conditions).

In order to give quantitative measures of the variability between results under these two extreme situations, precision is expressed in terms of repeatability and reproducibility. Hence many intermediate conditions are conceivable (for example day-to-day variations, between-instrument or operator variations within the same laboratory); repeatability and reproducibility have been found to deal with most practical cases.

4.2 Repeatability and reproducibility limits

In practice:

— two single results obtained within a laboratory under repeatability conditions should be considered suspect if they differ by more than r; and

— two single results obtained by two laboratories under reproducibility conditions should be considered suspect if they differ by more than R.

4.3 Systematic error or bias

Sources of systematic error or bias for indirect methods, and especially instrumental methods for milk analysis, may arise from error in calibration and variation in the chemical/physical form of the measured component, or from the influence of interfering factors. This last source of error is called “matrix effect”.

Therefore, according to the origin of the error and the ability of eliminating one source of error by adjustment of the calibration, the accuracy of the mean is split into two components: exactness of calibration (3.2.5) and accuracy of estimate or accuracy (3.2.6).

Accuracy measures the part of the systematic error not due to error in calibration.

Separating the bias into that due to calibration and that due to other factors serves two useful purposes: first, to enable a valid comparison of indirect methods, and second, to give a precise figure for the analytical performance requirements of instruments, and especially for the tolerances against the reference method.

When the adjustment of the calibration is the full responsibility of the operator, the exactness of calibration is not considered a performance characteristic of an indirect method.

However, under practical conditions, it is important to know exactly the actual accuracy of the mean of the method, that is, the extent of the average bias and the degree of uncertainty of individual results, whatever the origin of the systematic error may be. This question is considered in ISO 8196-2 as an application of precision and accuracy data.

5 Mathematical expressions

5.1 General

Mathematical expressions are derived from the analysis of variance for data obtained through an interlaboratory trial (see 6.1), the repeatability and reproducibility of a method are expressed for a given range of concentrations of the analyte by:

— the standard deviation of repeatability σ_r; and

— the standard deviation of reproducibility σ_R.

Use of the coefficient of variation (CV = σ/mean $\times$ 100) is recommended whenever the standard deviation varies proportionally with the level of the determinant.

This concept applies mainly to indirect methods calibrated against a reference method or standard materials, or when the true value of the component concentration is known.

5.2 Accuracy of the mean

This is calculated from the algebraic differences (d_i) between the measured and true values for a population of samples, and is expressed by:

— the mean of differences or bias ($\bar{d}$);

— the standard deviation of differences (σ_d);

— exactness of calibration: assuming that the relationship between the reference method values ($\bar{y}_i$) and the indirect method results ($\bar{x}_i$) is linear, the actual calibration of an instrument is considered exact if the slope of the linear regression equation ($\bar{y}_i = b\bar{x}_i + a$) is equal to 1,000 and the intercept equal to zero; in practice, the slope should not differ statistically from 1,000, nor the mean of the instrumental values $\bar{x}$ from the mean of the reference values $\bar{y}$ [mean of the differences $d_i = (x_i - y_i)$ not different from zero];

— accuracy: the accuracy of an indirect method is given by the critical values or tolerance limits within which the true value is estimated with a given probability, when the indirect method is exactly calibrated.

Accuracy is expressed by the residual standard deviation ($\sigma_{y,x}$) of the differences between the true value ($\bar{y}_i$) and the estimated mean reference value ($\hat{y}$) obtained from the regression of the actual calibration function (see Figure 1). It reflects, in a mathematical expression, the selectivity and specificity of the method in predicting the true values of the component to be measured.

In practice, for any individual sample, the mean $\bar{x}_i$ of test results given by an indirect method exactly calibrated should be considered suspect if the difference between $\bar{x}_i$ and the reference value $\bar{y}_i$ is outside the statistical tolerance limits.

6 Assessment of precision and accuracy

NOTE It should be borne in mind that both indirect and reference methods should be standardized or, at least, very accurate detailed procedures of the methods should be available before undertaking an overall accuracy evaluation.

6.1 Precision: Interlaboratory trial

6.1.1 General

For details concerning the organization and the statistical analysis, refer to ISO 5725-1.

In order to determine the precision (repeatability and reproducibility) of a test method, an interlaboratory trial or collaborative study is organized in which sub-samples of a certain number of samples covering the normal range of variation of the determinant are analysed in replicate by several laboratories using the same standard method. Variation of results, both within and between laboratories, enables the precision of the method to be estimated.

6.1.2 Experimental design

Preparation and performance of such a trial should be conducted under precise conditions, briefly summarized below.

a) The number of participating laboratories should be as large as possible ($p \geqslant 8$). In practice, with high-cost instrumental methods, it may be difficult to find equipment in numerous laboratories.

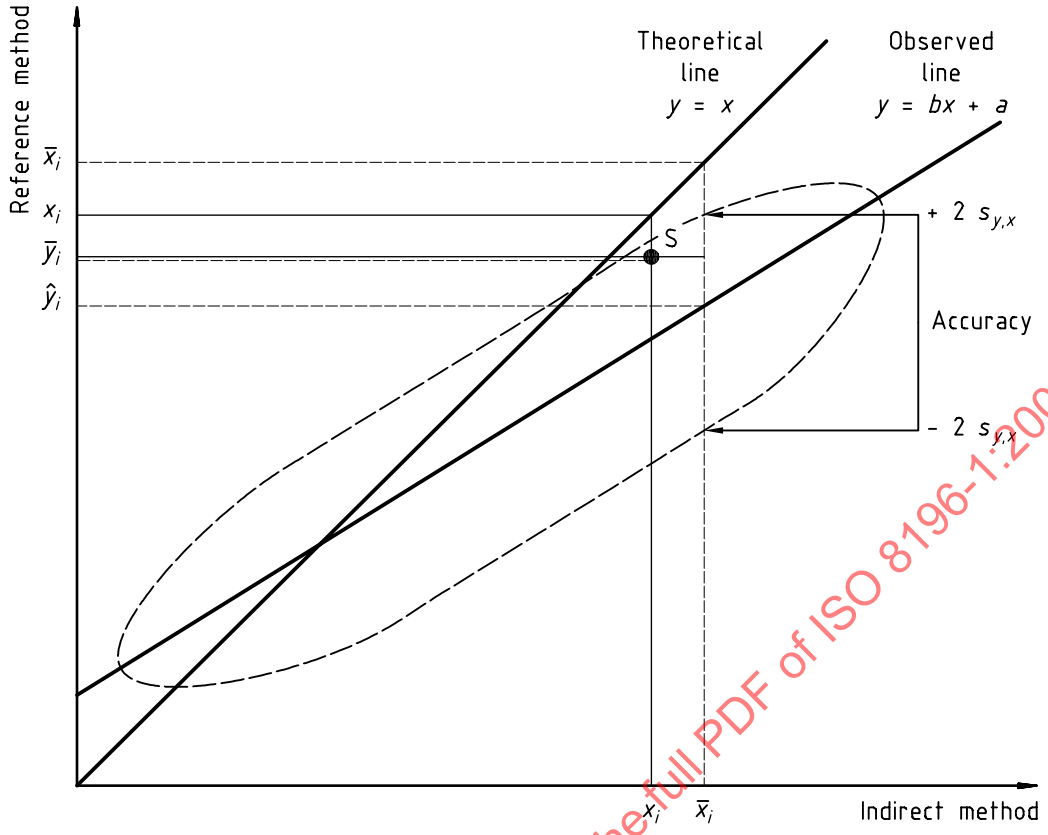
b) The method should be standardized or at least adequate instructions should have been written.

c) Assessing the reproducibility of an instrumental method implies that instruments are identically calibrated, otherwise reproducibility will reflect the variations in calibration rather than the real reproducibility of the instrument. One major source of difference between calibrations is the lack of uniformity of test results obtained with the reference method.

   To overcome this difficulty and obtain a valid estimate of reproducibility, the laboratory responsible for the collaborative study should distribute to each participating laboratory standard materials (e.g. milk samples) for calibration of the different instruments.

d) Milk samples should be as homogeneous as possible and no physical or chemical change in the samples should occur before analysis.

e) Laboratories should perform at least duplicate tests under repeatability conditions.



Key

x_i is the single instrumental value

$\bar{x}_i$ is the arithmetic mean of several determinations of sample S with the instrument

$\hat{y}_i$ is the estimated mean reference value for all the samples with an instrument level $\bar{x}_i$

$\bar{y}_i$ is the "true value" of the component to be measured for sample S

$s_{y,x}$ is the standard deviation from the regression

The mathematical model of the components of the total error on x_i is:

$$\underbrace{(x_i - \bar{y}_i)}_{\text{overall accuracy}} = \underbrace{(x_i - \bar{x}_i)}_{\text{repeatability}} + \underbrace{(\bar{x}_i - \bar{y}_i)}_{\substack{\text{accuracy of} \\ \text{the mean}}} + \underbrace{(\bar{x}_i - \hat{y}_i)}_{\substack{\text{exactness} \\ \text{of calibration}}} + \underbrace{(\hat{y}_i - \bar{y}_i)}_{\text{accuracy}}$$

Figure 1 — Breakdown of criteria in the overall accuracy of an indirect method

6.1.3 Statistical analysis

The fullest possible model of analysis of variance is given in this subclause.

This model is suitable providing that:

— within-laboratory (error) variances are homogeneous and there is no outlier result (see relevant part of ISO 5725 for specific statistical tests);

— r and R are independent of the level; if not, use a one-way analysis of variance for each level, or make appropriate data transformation (see 6.1.4);

— the laboratory bias may vary with the level, leading to a significant interaction (laboratory × level) effect;

— laboratories involved are considered as a random selection of the total laboratory population; this variable is assumed to be normal with a variance σ_L^2, called between-laboratory variance; this includes random and systematic differences between laboratories; usually, systematic differences which are unavoidable should be small;

— the component levels are also considered as a random selection of the total population of sample levels with a variance σ_S^2; this variance does not enter into the calculation of precision data.

Under these conditions a two-way analysis of variance using a cross-classification and a random model allows the calculation of the variance of repeatability σ_r^2 and the variance of reproducibility σ_R^2 (see Table 1).

Table 1 — Analysis of variance of precision data

Source of variation	Sum of squares	Degrees of freedom	Mean square	Expected mean square
Sample (level)	SOS_S	$q - 1$	$SOS_S/q - 1$	$s_e^2 + n \cdot s_{LS}^2 + n \cdot p \cdot s_S^2$
Laboratory	SOS_L	$p - 1$	$SOS_L/p - 1$	$s_e^2 + n \cdot s_{LS}^2 + n \cdot q \cdot s_L^2$
Laboratory × sample	SOS_{LS}	$(q - 1)\ (p - 1)$	$SOS_{LS}/(q - 1)\ (p - 1)$	$s_e^2 + n \cdot s_{LS}^2$
Error	SOS_e	$p \cdot q\ (n - 1)$	$SOS_e/p \cdot q\ (n - 1)$	s_e^2
NOTE Samples representing q different levels are sent to p laboratories which perform n replicate tests at each level.				

From Table 1 one can derive the following:

— s_r^2, since the estimate s_e^2 of the variance of error σ_e^2 is an unbiased estimate of σ_r^2;

— s_R^2, which is the sum of the variance between laboratories (s_L^2), the variance of interaction laboratory × sample (s_{LS}^2) and the variance of repeatability (s_r^2).

Thus: $s_R^2 = s_L^2 + s_{LS}^2 + s_r^2$

Finally, $r = 2{,}83 \times s_r$ and $R = 2{,}83 \times s_R$ are the quantitative expressions of the repeatability and the reproducibility limits of the method.

For a simple case for repeatability: when a set of q milk samples is analysed in duplicate by one laboratory, w_i being the absolute difference between duplicates, the standard deviation of repeatability can be calculated using the formula:

$$s_r = \left(\frac{1}{2q} \sum_{i=1}^{q} w_i^2 \right)^{\frac{1}{2}}$$