INTERNATIONAL STANDARD

ISO 8196-3

IDF 128-3

First edition
2009-10-01

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 pour l'évaluation et la validation des méthodes quantitatives alternatives d'analyse du lait

Reference numbers
ISO 8196-3:2009(E)
IDF 128-3:2009(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.org
Web www.iso.org

International Dairy Federation
Diamant Building • Boulevard Auguste Reyers 80 • B-1030 Brussels
Tel. + 32 2 733 98 88
Fax + 32 2 733 04 13
E-mail info@fil-idf.org
Web www.fil-idf.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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. 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 document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 8196-3|IDF 128-3 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.

ISO 8196|IDF 128 consists of the following parts, under the general title *Milk — Definition and evaluation of the overall accuracy of alternative methods of milk analysis*:

— *Part 1: Analytical attributes of alternative methods*

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

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

Foreword

IDF (the International Dairy Federation) is a non-profit organization representing the dairy sector worldwide. IDF membership comprises National Committees in every member country as well as regional dairy associations having signed a formal agreement on cooperation with IDF. All members of IDF have the right to be represented at the IDF Standing Committees carrying out the technical work. IDF collaborates with ISO in the development of standard methods of analysis and sampling for milk and milk products.

The main task of Standing Committees is to prepare International Standards. Draft International Standards adopted by the Action Teams and Standing Committees are circulated to the National Committees for voting. Publication as an International Standard requires approval by at least 50 % of IDF National Committees casting a vote.

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.

ISO 8196-3|IDF 128-3 was prepared by the International Dairy Federation (IDF) and Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 5, *Milk and milk products*. It is being published jointly by ISO and IDF.

All work was carried out by the Joint ISO-IDF Action Team on *Automated methods* of the Standing Committee on *Quality assurance, statistics of analytical data and sampling* under the aegis of its project leader, Mr. O. Leray (FR).

This edition of ISO 8196-3|IDF 128-3, together with ISO 8196-1|IDF 128-1 and ISO 8196-2|IDF 128-2, cancels and replaces IDF 128:1985, which has been technically revised.

ISO 8196|IDF 128 consists of the following parts, under the general title *Milk — Definition and evaluation of the overall accuracy of alternative methods of milk analysis*:

— *Part 1: Analytical attributes of alternative methods*

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

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

Introduction

This part of ISO 8196|IDF 128 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 too small a 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 part of ISO 8196|IDF 128 specifies a harmonized protocol for such a method validation by an expert laboratory. It lists the evaluation steps, 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, only a limited number of evaluations should suffice for a decision on approval either by national bodies or by an international organization for the application of the methods and/or equipment in their area. An example is given for the evaluation of a method for the determination of fat, protein, lactose, urea and somatic cell count 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 part of ISO 8196|IDF 128 specifies a protocol for the evaluation and validation of alternative quantitative methods of milk analysis.

The protocol is applicable to all milk components including somatic cells. For microbiological parameters other standards, such as ISO 16140[5], apply. This part of ISO 8196|IDF 128 is also applicable to the validation of new alternative methods where a limited number of analysts does not allow the organization of an interlaboratory study and ISO 8196-1|IDF 128-1, therefore, does not apply.

This part of ISO 8196|IDF 128 also establishes general principles of a procedure for granting international approvals of these alternative methods. These principles are based on the validation protocol defined in this part of ISO 8196|IDF 128.

2 Normative references

The following referenced documents are indispensable for the application 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 9622, *Whole milk — Determination of milkfat, protein and lactose content — Guidance on the operation of mid-infrared instruments*[1)]

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

1) Equivalent to IDF 141.

3 Terms and definitions

For the purpose of this document, the terms and definitions given in ISO 8196-1|IDF 128-1, ISO 8196-2|IDF 128-2, ISO 3534-1 and ISO 5725-1 apply, together with the following.

3.1
validation of an alternative method
demonstration that results obtained with an alternative method are comparable to those obtained with the reference method, thereby showing compliance of accuracy with defined requirements and fitness for purpose

3.2
measurand
component
analyte
criterion
particular quantity or characteristic subject to measurement

EXAMPLES A measurand may be a milk component, a physical characteristic or a biological element.

NOTE Adapted from ISO/IEC Guide 99:2007[8], 2.6.

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 organizing laboratory of an alternative method against the reference method under test bed conditions

3.5
method confirmation study
study performed in routine laboratories, of an alternative method to confirm results of a previous **methods comparison study** (3.4)

3.6
interlaboratory study
study of 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)

3.7
organizing laboratory
laboratory having qualified staff and equipment to perform a **methods comparison study** (3.4)

3.8
national approval
authorization of 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 official body

3.9
international approval
authorization of use of a method for defined purposes at the international level — generally for reasons of collective interest and/or having an official character — delivered by an international organization 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 respectively.

4.1.2 Phase I

A methods comparison study includes the assessment of the analytical attributes and a comparison of the alternative method against the reference method under test bed conditions. This part of the evaluation has to be carried out by an organizing laboratory specialized in analytical evaluations as well as being experienced in the application of the relevant reference method. The laboratory shall conform to ISO/IEC 17025 for this activity.

4.1.3 Phase II

A method confirmation study under routine testing conditions is initiated after a successful Phase I. The examination is recommended of at least two instruments located in different routine laboratories under routine testing conditions for a minimum period of two months. Care should be taken that each instrument is exposed to the level of sample variation normally expected during that period. Each instrument should fulfil the day-to-day quality control demands specified in ISO 8196-2|IDF 128-2 by checking compliance of results with figures of overall accuracy obtained in Phase I. The alternative method should also be assessed for general convenience aspects such as speed, consumables, user-friendliness, security, and robustness.

4.1.4 National approval

Based on the content of submitted reports, national bodies can authorize the use of an alternative method for defined purposes. Compliance with requirements stated in this protocol provides assurance of a sufficient quality in measurement results and comparability with other methods and/or instruments of a similar type validated elsewhere according to the same protocol.

4.1.5 International approval

International organizations can grant an international approval, e.g. for international milk recording, or to respond to a criteria approach. A number of successful individual (i.e. national) validations, reported in a standardized way, can provide sufficient confidence in the new alternative method performance and replace interlaboratory studies. The overall evaluation should be renewed successfully in a minimum number of distinct countries. Three independent validations are recommended.

4.2 Field of validity of the approval

4.2.1 An approval is given only under the circumstances specified in 4.2.2 and 4.2.3.

4.2.2 The field of application in which the instruments are used has been evaluated (component, concentration range, animal species, etc.). For instance, if milk of different animal species is to be analysed, specific evaluations for each species have to be carried out to assess that the instrument is appropriate for the expected use. If milk from breeds with unusual contents (e.g. Jersey breed with high fat and protein) is to be analysed, the evaluation should be carried out over the whole range of occurrence of the relevant component.

4.2.3 The specific method and/or instrument configuration used has been evaluated. If the configuration changes, proof should be obtained that it does not affect the precision and the accuracy beyond acceptable limits.

4.2.4 Carefully note and report all characteristics of both the milk products analysed 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 in Figure A.1. Each step corresponds to a source of error that may contribute to the overall uncertainty of the method (element in the breakdown of the overall accuracy). 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 first part of the protocol (5.2.2) is compulsory as it defines the minimum assessment sequence to be carried out.

A second part (5.2.3) is recommended to provide complementary information for future use.

5.2 Methods comparison study

5.2.1 General

This part specifies the elements of the evaluation which are compulsory.

The evaluation is to be carried out from test results expressed in standardized units of the reference method. For methods covering large ranges of measurand values (i.e. wider than 1 log unit), it is recommended to split the range into segments, each of maximum width one log unit, so as to obtain a minimum of three segments and to perform statistical calculations separately on each segment.

NOTE For instance, for fat in commercial milk, distinction can be made between skim milk, half-skim 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 (4.2); somatic cells in raw milk typically cover a range of several log units.

For methods where precision and accuracy are found to be proportional to the measurand value, apply an appropriate correction to the raw values.

Evaluation results should comply with 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 and somatic cells.

NOTE For liquid milk during milking or processing, there may be different assessment criteria for in-line and on-line analysis systems and at-line systems.

5.2.2 Compulsory assessments for the validation

5.2.2.1 Assessment of preliminary instrumental fittings

Before starting any further assessment, basic criteria indicating a proper functioning of the method or the instrument require verification. These criteria are daily precision (including repeatability and short-term stability), carry-over, and linearity.

5.2.2.1.1 Daily precision (repeatability and short-term stability)

Basically, the method used should present a measurement signal stability which complies with the precision requirements. If not, the analyser is either not functioning correctly (and should not be used) or its precision is not suitable for the objective of the analysis. Hence, the instantaneous stability (repeatability) and the signal level stability have to be assessed prior to any other characteristics.

EXAMPLE 1

The precision should be evaluated at three different concentration levels of the component measured: low, medium, and high. To achieve this, three different milk samples should be split into as many identical test portions as necessary for the analyses.

During the day, for each level, analyse the same milk sample in triplicate ($n = 3$) using the instrument every 15 min to 20 min without any change in the calibration in order to obtain a minimum of 20 check test series ($q \geqslant 20$). Preferably, it 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.

Using a one-way analysis of variance (ANOVA), estimate the standard deviation of repeatability, s_r, the standard deviation between check series, s_c, and the standard deviation of daily reproducibility, s_R, or, equivalently, according to the following:

For every check, j ($j = 1 \ldots q$), calculate the mean,

$$\bar{x}_j = \sum x_{ij}/n$$

and the standard deviation,

$$s_{rj} = \left[\sum (x_{ij} - \bar{x}_j)^2/(n-1)\right]^{1/2}$$

of replicates.

For the whole check sequence, calculate:

a) the standard deviation of repeatability: $s_r = \left(\sum s_{rj}^2/q\right)^{1/2}$

b) the standard deviation of means:

$$s_{\bar{x}} = \left[\sum (\bar{x}_j - \bar{x})^2/(n-1)\right]^{1/2} = \left\{\left[\sum \bar{x}_j^2 - \left(\sum \bar{x}_j\right)^2/q\right]/(q-1)\right\}^{1/2}$$

with

$$\bar{x} = \sum \bar{x}_j/q$$

c) the standard deviation between checks: $s_c = (s_{\bar{x}}^2 - s_r^2/n)^{1/2}$ with $s_c = 0$ if $s_c < 0$

d) the standard deviation of daily reproducibility: $s_R = (s_c^2 + s_r^2)^{1/2}$

EXAMPLE 2 The values obtained for s_r and s_R should comply with the limits stated in Annex B.

The stability of the method response during the sequence of check tests can be visualized by plotting measurement results and means, y, versus the check sequence numbers, x.

The significance of a possible observed deviation or fluctuation can be verified with the F-test of a one-way ANOVA or, equivalently, by calculating the observed value of F, F_{obs}:

$$F_{obs} = ns_{\bar{x}}^2/s_r^2$$

The test is significant if $F_{obs} > F_{1-\alpha}$ with $k_1 = q - 1$, $k_2 = q(n - 1)$, and $\alpha = 0{,}05$.

5.2.2.1.2 Carry-over effect

5.2.2.1.2.1 Strong differences in component concentrations between two successively analysed samples may 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 has to be clearly distinguished from rinsing efficiency.

5.2.2.1.2.2 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.

NOTE 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 $L_C \leqslant (r/\Delta L_{range}) \times 100$ 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 %.

5.2.2.1.2.3 The rinsing efficiency of the flow system has to be assessed separately by running tests without any correction (correction factor set to zero) in manual mode that bypasses the automated stirrer. Rinsing efficiency should not be less than 99 % or the internal carry-over should not exceed 1 %.

5.2.2.1.2.4 Analyse two samples, with high and low concentrations, respectively, of prior distribution in series of test portions. Repeat, as many times, N_C, 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_{L_1}, L_{L_2}, L_{H_1}, L_{H_2}. The minimum number of sequence replications, N_C, should be 20.

A sufficient number is recommended to reduce the relative uncertainty of the COR estimate, δ_{rel}, and to enable a clear differentiation from zero. A relative uncertainty of 20 % or less is sought. The relevant number of sequences can be obtained by $N_C \geqslant (100/\delta_{rel})^2$. Increasing the number of sequences is especially to be considered in case of estimating COR for adjustment of a correction factor.

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 $N_C \geqslant [r \times 100/(L_C \Delta L_{test})]^2$ where ΔL_{test} is the range between high and low concentration samples (equal to or greater than ΔL_{range}).

5.2.2.1.2.5 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.

Individual component concentrations have to 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.

NOTE 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 $\Delta L_{test} \geqslant r \times 100/(L_C \sqrt{N_C})$ 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. a 3 $\log_{10}$ scale or more, 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 are needed that can be extended to three for particularly wide ranges.

EXAMPLE For somatic cell counting in individual animal milk, the definition of three levels, at about 500×10^3 cells/ml, $1\,000 \times 10^3$ cells/ml, and $1\,500 \times 10^3$ cells/ml, is recommended.

5.2.2.1.2.6 Calculation: Calculate the mean and the standard deviations of the differences, $d_{L_L i} = L_{L_1 i} - L_{L_2 i}$ and $d_{L_H i} = L_{H_2 i} - L_{H_1 i}$, respectively, $\bar{d}_{L_L}$, s_{L_L}, $\bar{d}_{L_H}$, s_{L_H} and the mean difference of concentration, $\bar{d}_\rho = \bar{L}_{H_2} - \bar{L}_{L_2}$.

Then calculate the CORs, C, and their standard deviations, s_C, by using the following equations:

$$C_{H/L} = \bar{d}_{L_L} \times 100 / \bar{d}_\rho \text{ and } s_{C_{H/L}} = sd_{L_L} \times 100 / \bar{d}_\rho \sqrt{N_C}$$

$$C_{L/H} = \bar{d}_{L_H} \times 100 / \bar{d}_\rho \text{ and } s_{C_{L/H})} = sd_{L_H} \times 100 / \bar{d}_\rho \sqrt{N_C}$$

The COR can also be obtained by using the following equivalent formulas:

$$C_{H/L} = (\sum L_{L_1} - \sum L_{L_2}) \times 100 / (\sum L_{H_2} - \sum L_{L_2}) = (\bar{L}_{L_1} - \bar{L}_{L_2}) \times 100 / (\bar{L}_{H_2} - \bar{L}_{L_2})$$

$$C_{L/H} = (\sum L_{H_2} - \sum L_{H_1}) \times 100 / (\sum L_{H_2} - \sum L_{L_2}) = (\bar{L}_{H_2} - \bar{L}_{H_1}) \times 100 / (\bar{L}_{H_2} - \bar{L}_{L_2})$$

The two COR values obtained should not significantly differ from each other and should not exceed the limit, L_C, in the test condition stated for the component in Annex B.

Verify this by checking whether the following conditions are fulfilled:

$$C_{H/L} - C_{L/H} \geqslant t_{1-\alpha/2} \left[s^2_{C_{H/L}} + s^2_{C_{L/H}} \right]^{1/2}$$

$$C_{H/L} \leqslant L_C - t_{1-\alpha}\, s_{C_{H/L}}$$

$$C_{L/H} \leqslant L_C - t_{1-\alpha}\, s_{C_{L/H}}$$

with α = 0,05.

5.2.2.1.3 Linearity

5.2.2.1.3.1 **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.

NOTE Current alternative methods are frequently based on multiple signals using a multivariate approach. For these methods, in particular for examples involving small relative changes in the sample matrix and signals with low specificity, linearity assessment can be difficult due to large random error (low signal to noise ratio). In these cases, as the linearity error is contained in the overall accuracy component, linearity assessment can be omitted provided it is covered in the further step of accuracy evaluation.

The method is specified in 5.2.2.1.3.2 to 5.2.2.1.3.4.

5.2.2.1.3.2 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 and urea).

c) The linearity assessment range should be congruent with the concentration range for the validation study (Annex B).

d) Reference values for linearity samples can be established from either the mixing ratio or the theoretical concentrations as calculated from the concentrations of the initial samples. Depending on the alternative method, they should be obtained from volume by volume mixing ratios where analysis is performed on a milk volume (volumetric intake measurement) and mass by mass mixing ratios where analysis is applied to a weighed milk portion (see Annex E).

5.2.2.1.3.3 Analyses. Analyse each sample, firstly in order of increasing concentrations in $N_L/2$ replicates, secondly in order of decreasing concentrations in $N_L/2$ replicates, so as to obtain the total replicate number relevant for the measurand (see Annex B).

5.2.2.1.3.4 Calculation and assessment. Calculate the linear regression equation $y = bx + a$ ($y =$ instrument, $x =$ reference) and the residuals e_i ($e_i = y_i - bx_i - a$) from the means of replicates and the theoretical reference.

Plot the residuals, e_i, on the ordinate against theoretical concentrations on the abscissa. Visual inspection of the data points usually yields sufficient information about the linearity of the signal.

Any deviation from linearity or obvious trend in the data in this plot indicates a potential problem and should lead to further investigation of the method, as detailed below.

Any residual obviously being out of the current distribution (outlier) should lead to deletion of that result and repetition of the calculation before applying further tests.

Calculate the relative linearity bias by the ratio of the residual range to the signal values range:

$$\frac{\Delta e}{\Delta \rho} = \frac{e_{\max} - e_{\min}}{\rho_{\max} - \rho_{\min}}$$

where

$e_{\max}$ is numerical value of the upper residual;

$e_{\min}$ is the numerical value of the lower residual;

$\rho_{\max}$ is the numerical upper value measured with the instrument;

$\rho_{\min}$ is the numerical lower value measured with the instrument.

NOTE 1 Limits are defined from the prerequisite that deviation from linearity should not produce a larger error than the repeatability of the method over the usual measuring range. Hence, limits of the relative linearity bias, $L_{\Delta e/\Delta L}$, are meant to fulfil the condition $L_{\Delta e/\Delta L} \leqslant r/\Delta L_{\text{range}}$ for the upper acceptable repeatability, with r being the repeatability limit and ΔL_{range} being the difference between the maximum and the minimum concentration in the concentration range of interest. For components where repeatability is not constant over the measuring range, the relative linearity bias limits are set based on the levels of largest repeatability (e.g. somatic cell counting). Common limits for $\Delta e/\Delta L_{\text{range}}$ are in the range 0,01 to 0,02.

NOTE 2 The number of replicates needed to ensure significance of the $\Delta e/\Delta L$ test can be estimated by the conditions: $N_L \geqslant 8\,\sigma_r^2/(L_{\Delta e/\Delta L}^2\,\Delta L_{\text{test}}^2)$ or $N_L \geqslant r^2/(L_{\Delta e/\Delta L}^2\,\Delta L_{\text{test}}^2)$.

NOTE 3 Concentration ranges, ΔL_{test}, larger than ΔL_{range} allow the measurement of larger linearity bias, Δe, with a similar relative linearity bias and increased significance for the same maximum repeatability value. The minimum concentration range can be estimated by the conditions: $\Delta L_{\text{test}} \geqslant 2\sqrt{2}\sigma_r/(L_{\Delta e/\Delta L}\sqrt{N_L})$ or $\Delta L_{\text{test}} \geqslant r/(L_{\Delta e/\Delta L}\sqrt{N_L})$.

A one-way ANOVA can be carried out to confirm the statistical significance of non-linearity. Statistical tests for comparison of variances can be applied to confirm the significance of difference between residual variances.

Furthermore, if needed, non-linear trends can be approached by second and third degree polynomial and statistical tests, $\Delta e/\Delta L$ and F-tests used to select and assess the equation that allows the best linear fit.

Examples are given in Annex D.

5.2.2.1.4 Measurement limits

Limits of a measurement with an instrumental method exist at both extremities of the analytical range, e.g. a lower limit and an upper limit.

It is not required to determine these limits when natural concentration ranges for the respective components and species are normally located far from zero (which is generally the case for biochemical components, i.e. fat, protein, lactose, urea), and within the linearity range of the method.

The assessment of the measurement limits can be carried out in combination with the evaluation of the linearity. If linearity is not achieved throughout the whole concentration range, determine the actual range of application for the method concerned.

5.2.2.1.5 Lower limits

5.2.2.1.5.1 General. Lower limits are defined, as multiples of the standard deviation, σ, of random error observed near zero (blank), in three ways depending on the risk of error accepted and the precision requirements, as specified in 5.2.2.1.5.2 to 5.2.2.1.5.4.

5.2.2.1.5.2 Critical level, which is the smallest amount that can be detected (non-null) but not quantified as an exact value (risk $\beta = 50$ %). Below it cannot be assumed that the value is non-null:

$$L_{\text{crit}} = u_{1-\alpha}\,\sigma$$

EXAMPLES $L_{\text{crit}} = 1{,}645\sigma$ with $\alpha = 5$ %; $L_{\text{crit}} = 3\sigma$ with $\alpha = 0{,}13$ %.

5.2.2.1.5.3 Detection limit, for which the second type of error is minimized up to a defined level, generally equal to the level of risk, $\beta = 5$ %. It defines the lowest result, which differs significantly from zero (first type error, α), that can be produced with a sufficiently low probability (second type error, β) of including the blank value (zero) and with a sufficient confidence interval:

$$L_{\text{det}} = (u_{1-\alpha} + u_{1-\beta})\,\sigma$$

EXAMPLES $L_{\text{det}} = 3{,}29\sigma$ with $\alpha = \beta = 5$ %; $L_{\text{det}} = 6\sigma$ with $\alpha = \beta = 0{,}13$ %.

5.2.2.1.5.4 **Quantification limit**, or **determination limit**, which is the smallest amount of measurand that can be measured and quantified with a defined coefficient of variation (CV):

$$L_Q = k_q \, \sigma$$

with k_q = 100/CV since

$$CV = (\sigma / L_Q) \times 100 \quad \Rightarrow \quad L_Q = \sigma \times 100 / CV$$

EXAMPLES $CV = 10\ \% \ \Rightarrow \ L_Q = 10\sigma \ \Rightarrow \ \alpha = \beta < 0{,}000\ 1\ \%$

$CV = 30\ \% \ \Rightarrow \ L_Q = 3{,}3\sigma \ \Rightarrow \ \alpha = \beta = 5\ \%$

5.2.2.1.5.5 **Standard deviation**: Where, routinely, only single determinations are carried out, σ is the standard deviation of random error of the measurement. In the best case, that is the repeatability standard deviation at the proximity of zero content. Standard deviation of repeatability for the blank or standard deviation of repeatability estimated at concentrations close to zero are to be used. At least 20 samples are required as independent replicates.

EXAMPLE In automated somatic cell counting in milk, L_{det} may be expected to be not higher than 5 000 cells/ml. A measured σ value of 1 500 cells/ml indicates CV = 30 %. This value is to be compared to limits fit-for-purpose. However, the derived multiplying factor, k_q = 3,33, indicates shared risks of error, α and β, close to 5 % in accepting compliance with the limit 5 000 cells/ml.

In the absence of a total harmonization in definitions and multiplying factors, the major elements to assess are the standard deviation, σ, and the CV near the zero value. The risk associated with L_{det}, when compared to a specified limit value, needs to be accounted for (see example in C.1.4.2).

5.2.2.1.6 Upper limit

Upper limit corresponds to the threshold where the signal or the measurement deviates significantly from linearity (see 5.2.2.1.3).

Superseding the upper limit produces a ratio, $\Delta e/\Delta L$, exceeding accepted limits (see 5.2.2.1.3.4). Whether measured upper values deviate from linearity, i.e. whether y_U differs significantly from the linear prediction $y(x_U)$ without that result (see Annex D), can be checked:

$$t_{obs} = \left| y_U - y_U(x_U) \right| / sy(x_U)$$

with

$$sy(x_U) = s_{yx} \left[1 + (1/q) + (x_U - \bar{x})^2 / S_x \right]^{1/2}$$

If, with $q - 2$ degrees of freedom, and $\alpha = 0{,}05$:

$t_{obs} \leqslant t_{1-\alpha/2} \ \Rightarrow$ no deviation from linearity at that point;

$t_{obs} > t_{1-\alpha/2} \ \Rightarrow$ significant deviation from linearity at that point.

5.2.2.2 Evaluation of the overall accuracy

5.2.2.2.1 General

The overall accuracy is composed of the sum of the error of the repeatability, of the accuracy and of the calibration.

With raw milk, each part of the overall accuracy is measured through the analysis of individual milk samples and herd bulk milk samples of the animal species specified. Herd bulk milk samples shall be collected in addition to individual milk samples in order to measure more accurately that part of the variance related to herd effects.

The evaluation is to be performed under conditions equivalent to the intended operation in routine (working parameters, speed, and calibration).

NOTE Bulk milk and processed milk as commingled individual animal/herd bulk milk samples present an averaged matrix which results in very low accuracy error, close to the precision error, showing little variation in component concentrations. This makes them inappropriate for a proper evaluation of accuracy and calibration over a sufficient range. As precision is evaluated equivalently using animal and/or herd milk samples, the use of bulk milk samples is not required in the conditions stated in the present document.

5.2.2.2.2 Calibration

A preliminary but appropriate calibration (or pre-calibration) should be made as follows:

a) if a method or instrument is to be used without any further local calibration, alternative method analyses of the evaluation can be directly performed with appropriate (representative) milk samples;

b) if a local calibration is necessary, perform prior calibration in accordance with the manufacturer's instructions and equipment facilities, before starting up the evaluation. Prepare any calibration samples required in accordance with the specifications of the relevant International Standards for the measurand or, if no standardized procedure exists, in a similar way to the sample used in the evaluation.

5.2.2.2.3 Samples

Proper quality milk samples should be used. Individual milk samples should cover the maximum concentration range of the component, that is for the stated components according to the specifications of Annex B. Possible geographical, seasonal or regional variation in milk composition should be covered in the sample set.

The minimum number required for the measurand is related to the statistic significance in comparisons and the matrix variability of the product.

Whatever the milk, it cannot be lower than 50, thus providing a maximum 20 % relative uncertainty to standard deviation estimated from analytical results.

NOTE 1 The relative uncertainty, $\pm \delta_{\text{rel}}$, of standard deviation estimates can be obtained by: $\delta_{\text{rel}} \geqslant u_{1-\alpha/2} \times 100/\sqrt{2q}$, thus $q \geqslant 20\,000/\delta_{\text{rel}}^2$ for $q \geqslant 30$ samples and $\alpha = 0{,}05$.

For fat, protein and somatic cells in raw milk, generally a minimum number of 100 individual animal milk samples ($N_{\text{a}} \geqslant 100$) from different herds ($N_{\text{h}_1} \geqslant 5$) and 60 herd milk samples ($N_{\text{h}_2} \geqslant 60$) are required with regard to the need for sample representativeness (see Annex B).

NOTE 2 The sample number, q, should also allow the mean bias, $\bar{d}$, of calibration to be checked, thus minimizing the uncertainty lower than a defined limit, $L_{\bar{d}}$, that can ascertain evaluation requirements:

$$u_{1-\alpha/2}\sigma_{yx}/\sqrt{q} \leqslant L_{\bar{d}} \quad \Leftrightarrow \quad q \geqslant (2\sigma_{yx}/L_{\bar{d}})^2$$

or

$$u_{1-\alpha/2}\sigma_{yx,\text{rel}}/\sqrt{q} \leqslant L_{\bar{d},\text{rel}} \quad \Leftrightarrow \quad q \geqslant (2\sigma_{yx,\text{rel}}/L_{\bar{d},\text{rel}})^2$$

with $\alpha = 0{,}05$.

5.2.2.2.4 Assessment of repeatability

Repeatability is the main criterion indicating whether a method produces stable results according to user requirements. It is a major element of internal quality control. Therefore, every new instrument has to fulfil a maximum limit for repeatability value, stated in the relevant International Standard, in order to satisfy the criteria for approval.

Samples are to be analysed on the instrument calibrated according to the recommendations of the manufacturer, preferably in duplicate. Generally, series of 15 to 20 milk samples are successively analysed twice after recovering their initial analytical condition (i.e. temperature by heating) when necessary.

The standard deviation of repeatability is calculated from duplicate results obtained from the whole set of data and, for criteria covering a wide range of concentration, that is more than 1 log scale (in the case of somatic cell count), part by part after splitting of the whole concentration range into different parts (minimum three parts of maximum 1 log unit width each, i.e. low, medium, and high).

For q samples analysed in duplicate, the standard deviation of repeatability is calculated from (see also Annex D):

$$s_r = \left(\sum w_i^2 / 2q \right)^{1/2}$$

where w_i is the modulus of the difference between duplicates of sample i ($w_i = |x_{1i} - x_{2i}|$).

The value of s_r obtained should be compared with the limit value for the repeatability values, σ_r, as defined for the relevant measurand and application (see Annex B). It is expected that $s_r \leqslant \sigma_r$.

Additionally, to ascertain whether agreement with the limit is met with a probability, $p = 1 - \alpha$, the estimate s_r should also fulfil the condition $s_r \leqslant \sigma_r (X_\alpha^2/q)^{1/2}$ with $\alpha = 0{,}05$. It is up to the evaluator to choose the required degree of restrictiveness.

5.2.2.2.5 Assessment of trueness

5.2.2.2.5.1 General

According to ISO 8196-1|IDF 128-1, the error of trueness is broken down into the error of exactness of calibration and the error of accuracy (accuracy of estimates).

Statistical parameters to be used are indicated in ISO 8196-2|IDF 128-2:

$\bar{d}$, s_d, s_{yx}, slope, b, Student t-test for $\bar{d}$ and b, respectively.

The aforementioned parameters are obtained from a simple linear regression, calculated in accordance with ISO 8196-1|IDF 128-1 and ISO 8196-2|IDF 128-2 using means of duplicate instrumental results, x, and so-called reference results, y, obtained by the reference method in duplicate.

For measurands covering a wide concentration range, that is more than 1 log scale (i.e. with somatic cell count), accuracy evaluation should be performed for the whole range and for successive parts of the range after splitting the whole concentration range into different parts (at minimum three parts of maximum 1 log unit width each, i.e. low, medium, and high).

5.2.2.2.5.2 Assessment of accuracy

For raw milk, accuracy is assessed for individual animal milks and herd bulk milks separately. It is based on the residual standard deviation, s_{yx}, of the simple linear regression of instrumental results, x, and reference results, y. It is expected that the differences from the regression line are normally distributed. Any outlying results should be carefully scrutinized. For outlying results, further test samples drawn from the same calibration sample should preferably be reanalysed by both the reference method and the alternative method.

When outlying figures remain, the report should present s_{yx} estimates and graphs including all data — with the outliers identified, their number and respective biases — as well as s_{yx} estimates after discarding outliers. Statistical methods used to identify outliers should be specified in the evaluation report. The proportion of outliers should not exceed 5 %.

The estimated value of s_{yx} should fulfil the limits σ_{yx} as defined for the parameter and the matrix concerned. It should respect the condition $s_{yx} \leqslant \sigma_{yx}$. Limits for individual animal and herd bulk milk samples are given in Annex B.

Additionally, to ascertain fulfilment for not exceeding the limit with a probability $p = 1 - \alpha$, the estimate $s_{y,x}$ should also fulfil the condition $s_{yx} \leqslant \sigma_{yx}(X^2_{\alpha}/q)^{1/2}$ with $\alpha = 0{,}05$. It is up to the evaluator to choose the required degree of restrictiveness.

5.2.2.2.5.3 Assessment of exactness of calibration

Prior to analyses, calibrate the alternative method according to the procedure recommended by the manufacturer. Express the results in the same unit as those of the reference method. Raw signals and further statistical comparisons can then be made on the same scale, allowing statistical tests of conformity and assessments against prior stated target values. For this purpose, individual animal and herd bulk milk samples should be analysed to provide the relevant information on the quality of the adjustment.

Depending on the principle of the method, the quality of the calibration can be influenced by the representativeness of the calibration sample set. This is in addition to the influence of the calibration technique applied (i.e. mathematical model, experimental design, process). Therefore, the error due to non-representativeness shall be maximally reduced, e.g. by sampling calibration samples under conditions close or identical to those for prediction samples.

Assess the exactness of calibration according to ISO 8196-2|IDF 128-2 using the parameters of the regression, $y = bx + a$, with the mean bias, $\bar{d}$, and the slope, b, taking care of any outlying results and fitting an appropriate line through the dot distribution of results on a graph. Estimates $\bar{d}$ and b should normally fulfil the limits defined for the component and the product (e.g. in Annex B). Failing that goal normally implies further investigations or explanations.

NOTE 1 Low correlation of alternative method results with reference method results can turn into a poor line fitting on to the dot distribution. For methods showing small correlation (e.g. $r_{xy} < 0{,}90$) and in cases where the regression line obviously deviates from the axis of the dot distribution, pooling of samples can improve slope evaluation. It can be simulated mathematically by arithmetically averaging individual sample results for 10 to 15 successive groups of equal sample numbers along the range, after ranking according to increasing alternative method results. By this method, more accurate indication on line adjustment is obtained through reducing the residual standard deviation, whereas the range is only slightly diminished and representativeness is maintained. Further calculations are then carried out using the means obtained to assess the slope.

NOTE 2 Compliance limits in Annex B are generally larger than significance limits expected from usual statistical tests so as to take into account possible herd effects in individual milk samples on slope, and also uncertainty of both adjustment and assessment of calibration. Users can decide to be more demanding than these limits.

5.2.3 Additional informative investigations

5.2.3.1 General

The following items are not compulsory elements for an evaluation even though they are of interest as possible contributors to the overall accuracy of the method. Moreover, knowledge gained about the method may have implications in milk sample handling (sampling, preservation, shipment, etc).

5.2.3.2 Ruggedness

Ruggedness is the ability of a method not to be influenced by external elements other than the component measured itself. Possible effects can come from concentration variation of major milk components or interactions, biochemical changes of milk components related to preservation (lipolysis, proteolysis, lactic souring) or chemicals added to the milk such as preservatives.

The principle of ruggedness assessment is to produce a significant change in the concentration of each interacting component separately and measure the corresponding change in the measurement result of the influenced component.

Then the ratio of the difference observed and the change introduced is calculated and expressed in the relevant units.

5.2.3.2.1 Effect of major milk components (interference)

To determine interactions of milk components (fat, protein, lactose), appropriate guidance is provided in ISO 9622 for sample preparation and calculation procedures. Although at first developed for mid-infrared methods, the approach is applicable also to other methods.

As for lactose, the effect of urea on other component measurements can be evaluated by addition of urea to milk.

The effect of high fat and protein content on somatic cell count in milk (sheep, goat and buffalo milk) can be evaluated by recombining cream (natural creaming) and milk retentate in a similar way to that specified in ISO 9622.

The effect should preferably be measured at three relevant levels within the range of the measurand, i.e. low, medium, and high, for the animal species.

5.2.3.2.1.1 Effect of biochemical changes in components

Biological changes in milk usually result in breakdown of milk components that can be induced by bacterial growth or enzymatic activity. Deterioration of milk samples may go unnoticed. Therefore, it is relevant to check the susceptibility of an alternative method for such deterioration, in particular in order to evaluate the quality of the sample preservation and the suitability of sampling and shipment conditions.

Clotting, churning and oiling are generally clearly visible defects of raw milk that may affect analytical results. In those cases, samples should be discarded.

5.2.3.2.1.2 Lipolysis

The possible effect of lipolysis can be monitored through artificial induction (i.e. repeated cooling, heating, and vigorous mixing) or through activation of native lipase or through addition of bacterial lipase (e.g. from *Pseudomonas* spp.). One should raise the level of free fatty acids up to at least 5 meq/100 g of fat.

At least five levels are required. The effect exists if the slope of a linear regression equation of measurement result, y, versus free fatty acid concentration, x, is significantly different from 0,00.

5.2.3.2.1.3 Proteolysis

The possible effect of proteolysis can be monitored through artificial induction (i.e. using microflora proteases). A minimum range of 0,8 % soluble nitrogen in milk should be obtained.

At least five levels are required. The effect exists if the slope of a linear regression equation of measurement result, y, versus percentage mass fraction of soluble nitrogen, x, is significantly different from 0,00.

5.2.3.2.1.4 Lactic souring, pH

The possible effect of souring can be monitored through addition of lactic acid.

At least five levels are required. Check that at the higher levels the milk does not clot at the water-bath temperature in order to prevent blockage or damage to the liquid flow systems.

The effect exists if the slope of a linear regression equation of measurement result, y, versus lactic acid concentration, x, is significantly different from 0,00.

5.2.3.2.2 Effect of sample history and handling conditions

Combination of cooling and storage at 0 °C to 6 °C with a preservative such as bronopol (2-bromo-2-nitropropan-1,3-diol) appropriately preserves clean (uncontaminated) milk samples. These conditions generally apply to calibration and control milk samples. In practice, sample conditions may differ (different type of chemical preservatives, transport, storage time and temperatures).

Therefore, it is of interest to determine the effect of preservation conditions on alternative method results. Based on this, adequate advice on proper sample handling and preservation can be provided to workers involved.

This is achieved by analysing two identical sample sets with the method in one run, one set having the usual preservation and having undergone the usual handling conditions, the other set having been subject to optimal conditions, and then comparing the obtained results.

For each item, component concentrations should cover the usual concentration range in practice. Sample numbers of 30 to 40 are generally sufficient. Statistics to be used are the same as in the assessment of overall accuracy, using optimal conditions as reference. The mean and the standard deviation of differences provide information on the average error due to sample conditions tested and possible significance. Additionally, effect on repeatability can be evaluated by analysing duplicates and comparing s_r obtained in both preservation/handling conditions using the F-test.

5.2.3.2.2.1 Effect of added chemicals (preservatives)

A possible effect on analytical results can be monitored through comparisons of identical parallel series of milk samples preserved with different chemical preservatives. Other preservation parameters have to be maintained equal in order not to influence the results. Evaluate the effect of both the nature and concentration of the preservative.

5.2.3.2.2.2 Effect of sample intake temperature

Analytical instruments may be sensitive to environmental conditions (i.e. humidity, temperature, vibrations). In particular, sample temperature may be a critical point with respect to internal instrument temperature. A comparison at two extremes (specified lower and upper limit as advised by the manufacturer) on identical sets of different milk samples provides sufficient information.

5.2.3.2.2.3 Effect of storage conditions (time and temperature)

Sample temperature may affect the physical characteristics of milk components (i.e. crystallization of fat solubility of casein and the mineral fraction).

Besides, storage time can determine the ability of milk to recover its native physical and chemical characteristics before being analysed. For instance, cream separated from skim milk becomes so firm that obtaining a homogenous sample is hampered when applying normal mixing conditions.

5.2.3.3 Practical convenience (Phase II)

This part of the evaluation consists of various elements that determine the ability of the laboratory to produce analytical results within the time expected and at the cost expected or needed.

During the Phase II period, the instrument should be used continuously for at least 4 h per day. The total number of continuous work rounds should be a minimum of 20 in each of the laboratories designated for the evaluation under routine conditions.

For the following items, opinions should be provided on criteria such as compliance with expected performance, simplicity, reliability, quickness, and security, where relevant, using three mark levels or grades (poor, medium, good) and including specific remarks.

5.2.3.3.1 Testing rate

The testing capacity of the alternative method, as indicated by the manufacturer, has to be verified. Precision performance should be reported at different speed levels, if relevant.

5.2.3.3.2 Robustness

Frequency of instrumental or other errors and servicing operations required should be recorded with a description of the nature of the incidents.

5.2.3.3.3 User friendliness

Convenience for the user should be evaluated. Relevant items are:

a) ease of use of interfaces and software;

b) trouble-shooting facilities;

c) ease of servicing;

d) ease of calibration.

5.2.3.3.4 Reagents, waste and environmental aspects

Where relevant, mention should be made of local regulations concerning waste disposal, risks of environmental pollution, and health and safety of laboratory staff called upon to handle chemical or biological reagents and materials.

5.2.3.3.5 Recovery of precision in routine conditions

Via the application of internal quality control, results of quality checks, as recommended in ISO 8196-2|IDF 128-2, should be collected and reported in order to consolidate the results of Phase I. They can be summarized in tables, indicating for each type of check the successive check parameter values collected during the test period, their means, minima and maxima, total check numbers and degree of compliance.

5.3 Report and approval delivery

5.3.1 General

Data and experience gathered in both Phases I and II shall be duly reported in specific documents with all the necessary information on the evaluation course, tables with results on analytical performances, discussion, conclusions and summaries, and be made available for an eventual approval procedure.

5.3.2 National validation

The overall report should comprise four documents:

a) report of the test bed evaluation (Phase I), including raw results in annexes;

b) report of the first evaluation in routine conditions (Phase II);

c) report of the second evaluation in routine conditions (Phase II);

d) summary of conclusions of the three reports (Phases I and II) and a general conclusion in regard to the intended use of the method.

5.3.3 International validation

The reports of the national validations in three different countries (see definitions in Clause 3), are collected for transmission to the international validation body by the requesting organization with the respective validation certificates delivered by national bodies. The formal request should be made according to the procedure defined by the international organization granting the international validation or approval. It should comprise the elements proving prior successful validations in the required number.

Examples of this are:

a) the formal request for the validation/approval with appropriate forms where available;

b) the technical documentation relating to the method/instrument (i.e. principle, device and capabilities) as supplied by the manufacturer;

c) the national reports and national validation certificates from three different countries.

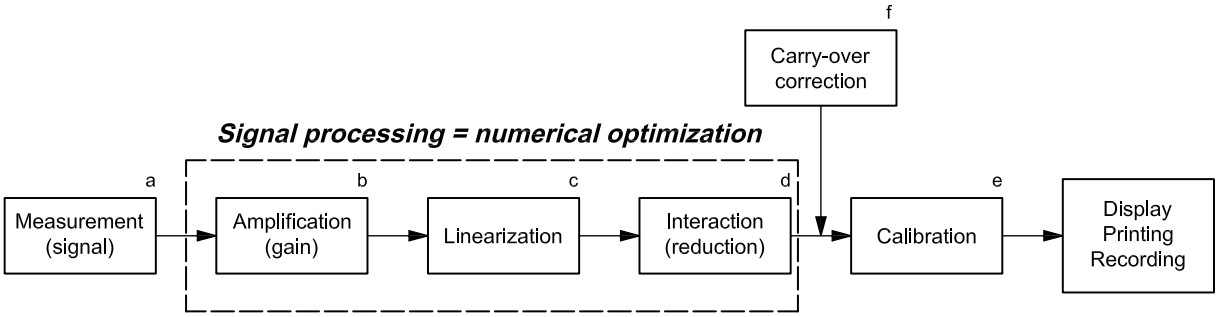
Annex A
(informative)

Measurement process and overall accuracy

Whatever the indirect method is, Figure A.1 represents the measurement process schematically. Not all steps necessarily exist in every method or instrument. This depends on manufacturer choice in relation to the principle of the measurement and the component measured. For instance, only a small or negligible effect of fat and protein is to be expected in the interaction (reduction) step in somatic cell counting by fluoro-opto-electronic methods in milk, as interactions are normally overcome by dispersion through treatment with chemical reagents before the measurement.

In some cases, several steps can be combined, e.g. those lying within the dashed line box in Figure A.1 for particular infrared devices. Nevertheless, in theory, the different steps of the signal processing can be set up in the instrument and remain available to be activated or not, through active or neutral mathematical matrices.

Interactions of major components or carry-over effects can be accommodated by adapting the principle of the method and/or the physical device (physical treatment, chemical reagents, tube length) and therefore no longer need numerical corrections.



Key

a Zero/blank, repeatability, stability, reproducibility.

b Sensitivity, measurement lower limit, repeatability.

c Linearity range, upper limit, accuracy.

d Effect of other milk components, accuracy.

e Suitability of manufacturer calibration system, accuracy.

f Effect of previous milk intake, repeatability, accuracy.

Every step of the measurement process corresponds to an element of the breakdown of overall accuracy of the method. Minimizing the overall error is achieved through minimizing every component thereby optimizing every step of the measurement process. Then the experimental design for the evaluation of a milk analyser is defined in order to assess that every measurement step is correctly adjusted.

Figure A.1 — Example of a theoretical measurement process in conventional analysers

Annex B
(informative)

Limits for the performance characteristics with raw milk

B.1 Limits for milk with medium fat and protein content

The limits listed in Table B.1 apply to cow and goat milk samples with medium content in the ranges mentioned.

Table B.1 — Limits with milk of medium fat and protein content

Measurand (units)	Criteria limits				
	Fat g/100 g	**Protein**[a] g/100 g	**Lactose** g/100 g	**Urea** mg/100 g	**SCC** 1 000 cells/ml
Range, ΔL_{range} — Whole	2,0 to 6,0	2,5 to 4,5	4,0 to 5,5	10,0 to 70,0	0 to 2 000
— Low (L)					0 to 100
— Medium (M)					100 to 1 000
— High (H)					>1 000
Carry-over ratio limit, L_C	1	1	1	2	2 [b]
Sequence number, N_C	20	20	20	20	20
Minimum range, ΔL_{test}	4	3	1,5	45	500
Linearity: ratio limit, $\Delta e/\Delta L$	0,01	0,01	0,01	0,02	0,02
Replicate number, N_L	6	6	6	6	8
Minimum range, ΔL_{test}	4	4	4	100	2 000
Repeatability					
Average standard deviation, s_r	0,014	0,014	0,014	1,4	
Coefficient of variation, $s_{r,rel}$ — Whole					4 %
— Low (L)					8 %
— Medium (M)					4 %
— High (H)					2 %
Within-lab. reproducibility					
Average standard deviation, s_R	0,028	0,028	0,028	2,8	
Coefficient of variation, $s_{R,rel}$ — Whole					5 %
— Low (L)					10 %
— Medium (M)					5 %
— High (H)					2,5 %
Accuracy					
Animal samples					
Standard deviation, s_{yx}	0,10	0,10	0,15	6,0	
Coefficient of variation, $s_{yx,rel}$					10 %
Animal number, N_a	100	100	100	100	100
Herd number, N_{h1}	5	5	5	5	5
Herd samples					
Standard deviation, s_{yx}	0,07	0,07	0,07	4,0	
Coefficient of variation, $s_{yx,rel}$					10 %
Herd number, N_{h2}	60	60	60	60	60
Calibration					
Mean bias, $\bar{d}$	± 0,05	± 0,05	± 0,05	± 1,2	
Relative mean bias, $\bar{d}_{rel}$					± 5 %
Slope, b	1 ± 0,05	1 ± 0,05	1 ± 0,10	1 ± 0,10	1 ± 0,05

[a] Same criteria limits for protein apply to crude protein, CP ($w_{CP} = w_{N,tot} \times 6,38$), true protein, TP [$w_{TP} = (w_{N,tot} - w_{N,non\text{-}P}) \times 6,38$] and casein, Cas [$w_{Cas} = (w_{N,tot} - w_{N,non\text{-}C}) \times 6,38$] where $w_{N,tot}$ is total nitrogen content; $w_{N,non\text{-}P}$ is non-protein nitrogen content; and $w_{N,non\text{-}C}$ is non-casein nitrogen content.

[b] Limit for $C_{H/L}$.

B.2 Limits for milk with high fat and protein content

The limits listed in Table B.2 apply to samples of sheep milk, buffalo milk and to milk of particular breeds of cow and goat with high fat and protein contents in the ranges mentioned.

Table B.2 — Limits for milk with high fat and protein content

Criteria (units)	Criteria limits				
	Fat g/100 g	**Protein**[a] g/100 g	**Lactose** g/100 g	**Urea** mg/100 g	**SCC** 1 000 cells/ml
Range, ΔL_{range} — Whole	5,0 to 14,0	4,0 to 7,0	4,0 to 5,5	10,0 to 70,0	0 to 2 000
— Low (L)					0 to 100
— Medium (M)					100 to 1 000
— High (H)					> 1 000
Carry-over ratio limit, L_C	1	1	1	2	2 [b]
Sequence number, N_C	20	20	20	20	20
Minimum range, ΔL_{test}	4	3	1,5	45	500
Linearity: ratio limit, $\Delta e/\Delta L$	0,01	0,01	0,01	0,02	0,02
Replicate number, N_L	6	6	6	6	8
Minimum range, ΔL_{test}	4	4	4	100	2 000
Repeatability					
Average standard deviation, s_r	0,028	0,028	0,014	1,4	
Coefficient of variation, $s_{r,rel}$ — Whole	0,35 %	0,40 %	0,30 %		4 %
— Low (L)					8 %
— Medium (M)					4 %
— High (H)					2 %
Within-lab. reproducibility					
Average standard deviation, s_R	0,056	0,056	0,028	2,8	
Coefficient of variation, $s_{R,rel}$ — Whole	0,70 %	0,80 %	0,60 %		5 %
— Low (L)					10 %
— Medium (M)					5 %
— High (H)					2,5 %
Accuracy					
Animal samples					
Standard deviation, s_{yx}	0,20	0,20	0,15	6,0	
Coefficient of variation, $s_{yx,rel}$	2,5 %	3,0 %			10 %
Animal number, N_a	100	100	100	100	100
Herd number, N_{h_1}	5	5	5	5	5
Herd samples					
Standard deviation, s_{yx}	0,14	0,14	0,07	4,0	
Coefficient of variation, $s_{yx,rel}$	1,75 %	2,0 %			10 %
Herd number, N_{h_2}	60	60	60	60	60
Calibration					
Mean bias, $\bar{d}$	± 0,10	± 0,10	± 0,05	± 1,2	
Relative mean bias, $\bar{d}_{rel}$	± 1,25 %	± 1,5 %			± 5 %
Slope, b	1 ± 0,05	1 ± 0,05	1 ± 0,10	1 ± 0,10	1 ± 0,05

[a] Same criteria limits for protein apply to crude protein, CP ($w_{CP} = w_{N,tot} \times 6{,}38$), true protein, TP [$w_{TP}$V($w_{N,tot} - w_{N,non\text{-}P}$) × 6,38] and casein, Cas [$w_{Cas} = (w_{N,tot} - w_{N,non\text{-}C}) \times 6{,}38$] where $w_{N,tot}$ is total nitrogen content; $w_{N,non\text{-}P}$ is non-protein nitrogen content; and $w_{N,non\text{-}C}$ is non-casein nitrogen content.

[b] Limit for $C_{H/L}$.

Annex C
(informative)

Calculation examples

C.1 Assessment of preliminary instrumental fittings

C.1.1 Daily precision

As an example, daily precision data from fat analysed by infrared spectroscopy (cf. ISO 9622) are listed in Table C.1.

Table C.1 — Daily precision

Check q	Replicate results x	Sum	Mean $\bar{x}\ \mu$	Mean bias $\bar{d}$	Test number n	Sum of squares S	Variance V	Standard deviation s_{ri}
1	4,00 4,03 4,01	12,04	4,013	0,008	3	0,000 467	0,000 233	0,015
2	4,02 4,03 4,02	12,07	4,023	0,018	3	0,000 067	0,000 033	0,006
3	4,01 4,00 4,00	12,01	4,003	–0,002	3	0,000 067	0,000 033	0,006
4	3,99 4,00 4,02	12,01	4,003	–0,002	3	0,000 467	0,000 233	0,015
5	3,99 4,01 4,01	12,01	4,003	–0,002	3	0,000 267	0,000 133	0,012
6	3,97 3,99 4,00	11,96	3,987	–0,018	3	0,000 467	0,000 233	0,015
7	4,01 4,00 3,98	11,99	3,997	–0,008	3	0,000 467	0,000 233	0,015
8	4,02 4,02 3,99	12,03	4,010	0,005	3	0,000 600	0,000 300	0,017
9	4,01 4,00 4,03	12,04	4,013	0,008	3	0,000 467	0,000 233	0,015
10	3,99 3,99 4,01	11,99	3,997	–0,008	3	0,000 267	0,000 133	0,012
Sum	120,150	120,150	40,050	0,000	30	0,003 60	0,001 80	
Average	4,005		4,005	0,000		0,000 180	0,000 180	0,013
Std deviation			0,010	0,010				

Check on homogeneity of variances using the Cochran test. In the following:

I_{Coch} is the Cochran index;

L_{Coch} is the Cochran limit:

L_s is the standard deviation limit;

$s_{\text{obs},i}$ are the standard deviation values observed;

$V_{\max}$ is the maximum variance;

ΣV_{i} is the sum of variances.

$$I_{\text{Coch}} = \frac{V_{\max}}{\sum V_i} < L_{\text{Coch}} \quad \Rightarrow \quad L_s = (L_{\text{Coch}} \sum V_i)^{1/2} \quad \Rightarrow \quad L_{\text{Coch}}(p = 0{,}95; 2; 10) = 0{,}445 \quad \Rightarrow$$
$$L_s = 0{,}028\,3 \geqslant s_{\text{obs},i}$$

The fact that the standard deviation limit is never smaller than the standard deviation values observed implies that the variance homogeneity is confirmed.

Daily reproducibility:

$$s_R = \left[s_{\bar{x}}^2 - s_r^2 \left(1 - 1/n\right) \right]^{1/2} \qquad s_R = 0{,}015 < 0{,}028$$

This implies that daily reproducibility is in conformity with ISO 9622.

Standard deviation between checks:

$$s_{\text{c}} = (s_{\bar{x}}^2 - s_r^2 / n)^{1/2} \qquad s_{\text{c}} = 0{,}007$$

Repeatability:

$$s_r = \left[\sum s_{r,i}^2 / q \right]^{1/2} \qquad s_r = 0{,}013 < 0{,}014$$

This implies that repeatability is in conformity with ISO 9622.

Source of variation	Degrees of freedom ν	Sum of squares S	Mean square S/ν	Standard deviation s	F
Intercheck	9	0,002 950	0,000 328	0,018	1,821
Intracheck	20	0,003 600	0,000 180	0,013	
Total	29	0,006 550	0,000 226	0,015	

Because $F_{\text{obs}} = 1{,}82$ is smaller than $F_{0{,}95} = 2{,}39$, it can be concluded that stability is assessed positively: no significant shift of instrument response observed.

Another conclusion is that, as the residual standard deviation, at 0,013, is smaller than 0,014, instrument functioning is assessed positively: no abnormal individual fluctuation.

C.1.2 Carry over effect

As an example, carry-over effect data from fat analysed by infrared spectroscopy (cf. ISO 9622) are listed in Table C.2.

Table C.2 — Carry-over effect

Sequence No.	Concentrations				Differences	
	L_{L_1}	L_{L_2}	L_{H_1}	L_{H_2}	ΔL_L	ΔL_H
1	0,00	–0,01	3,98	3,99	0,010	0,010
2	0,01	–0,01	3,99	4,01	0,020	0,020
3	0,00	–0,02	3,97	3,99	0,020	0,020
4	–0,01	–0,02	3,97	3,98	0,010	0,010
5	–0,01	–0,02	3,96	3,98	0,010	0,020
6	0,01	0,00	3,98	4,00	0,010	0,020
7	0,00	–0,02	3,99	4,01	0,020	0,020
8	0,01	–0,01	3,97	3,99	0,020	0,020
9	–0,01	–0,02	3,98	3,99	0,010	0,010
10	0,01	–0,01	3,99	4,00	0,020	0,010
Mean	0,001	–0,014	3,978	3,994	0,015	0,016
Standard deviation	0,009	0,007	0,010	0,011	0,005	0,005
Number, N	10	10	10	10	10	10
Student t	—	—	—	—	9,00	9,80
Minimum	–0,01	–0,02	3,96	3,98	0,01	0,01
Maximum	0,01	0,00	3,99	4,01	0,02	0,02
Δ = max. – min.	0,02	0,02	0,03	0,03	0,01	0,01

Mean bias $\overline{d}_{L_L}$ and $\overline{d}_{L_H}$ are significant according to the Student t-test $t_{0,975} = 2,26$

	Value	Confidence limit		Conclusion
		Lower	Upper	
$C_{L/H}$ =	0,40	0,31	0,49	COR lower than 1 % ⇒ conformity
$C_{H/L}$ =	0,37	0,28	0,47	COR lower than 1 % ⇒ conformity

C.1.3 Assessment of linearity

C.1.3.1 As a first example, linearity data from fat analysed by infra red spectroscopy (cf. ISO 9622) are listed in Table C.3.

Table C.3 — Linearity — Test sample set with progressive dilution of 10 % fat milk by skim milk

Level No.	Dilution mass per volume x %	Replicates 1	2	3	Mean concentration $\bar{y}$	Mean residual $\bar{e}$	Standard deviation s_r
1	15,50	1,54	1,52	1,53	1,530	–0,023	0,010
2	20,35	2,02	2,02	2,02	2,020	–0,013	0,000
3	25,64	2,55	2,56	2,55	2,553	–0,003	0,006
4	31,18	3,10	3,11	3,12	3,110	0,005	0,010
5	34,80	3,49	3,48	3,49	3,487	0,024	0,006
6	39,80	3,97	3,99	4,00	3,987	0,029	0,015
7	45,15	4,50	4,50	4,51	4,503	0,016	0,006
8	50,50	5,02	5,02	5,01	5,017	0,000	0,006
9	56,65	5,61	5,63	5,62	5,620	–0,006	0,010
10	61,95	6,11	6,13	6,12	6,120	–0,030	0,010
Number, N	10,0	10,0	10,0	10,0	10,0	10,0	
Mean	38,152	3,791	3,796	3,797	3,795	0,000	
Standard deviation	16,428	1,622	1,631	1,626	1,626	0,020	0,009
Minimum	15,500	1,540	1,520	1,530	1,530	–0,030	
Maximum	61,950	6,110	6,130	6,120	6,120	0,029	
Δ = max. – min.	46,450	4,570	4,610	4,590	4,590	0,059	

Linear regression on:	Replicates	Means
slope	0,098 98	0,098 98
intercept	0,018 56	0,018 56
data number, N	30	10

Standard deviation of residual means: $s_e = 0{,}020\ 3$

Standard deviation of repeatability: $s_r = 0{,}008\ 8$

Standard deviation of level biases: $s_L = 0{,}019\ 7$

from $s_L = (s_e^2 - s_r^2 / n)^{1/2}$

On performing a statistical test to evaluate the $\Delta e / \Delta L$ ratio, a Δe of 0,059 and ΔL of 4,590 give a value of 0,013. As this is greater than 0,01, this implies that linearity is inadequate.

A second statistical test to evaluate bias from the linearity test using the standard deviation of residual means was performed. The value of F_{obs}, given by:

$$F_{obs} = \left(s_r^2 + n s_L^2\right) / s_r^2 = n s_e^2 / s_r^2$$

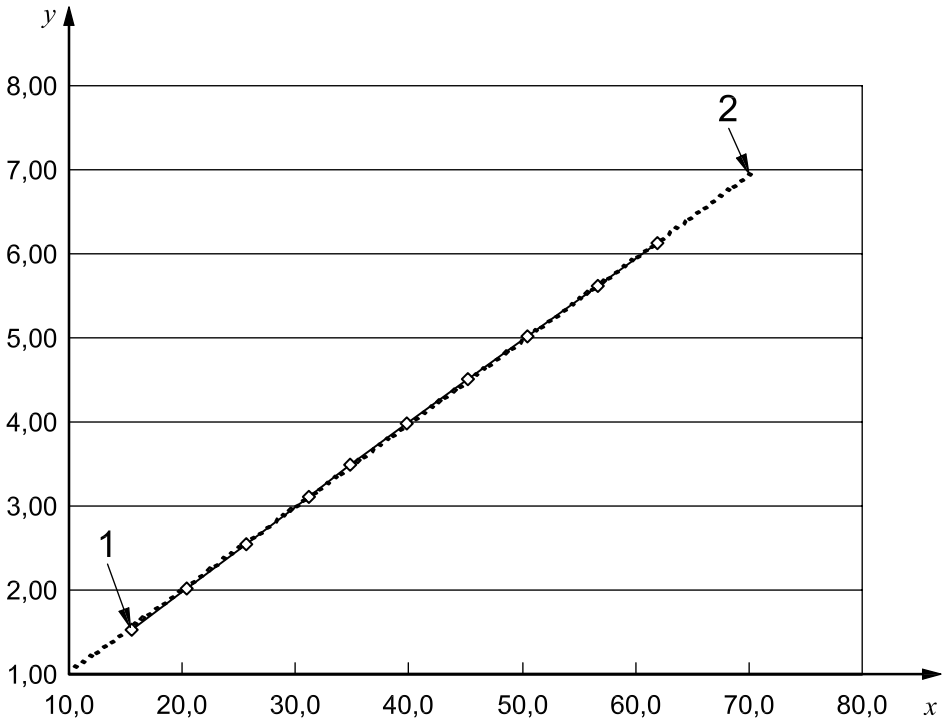
should be lower than $F_{0,95} = 2{,}45$ with $k_1 = q - 2$ and $k_2 = q(n - 1)$ degrees of freedom.

With $k_1 = 8$ and $k_2 = 20$

$F_{obs} = 16{,}17 > F_{0,95} = 2{,}45$

which implies that linearity is inadequate.

See Figures C.1 and C.2.

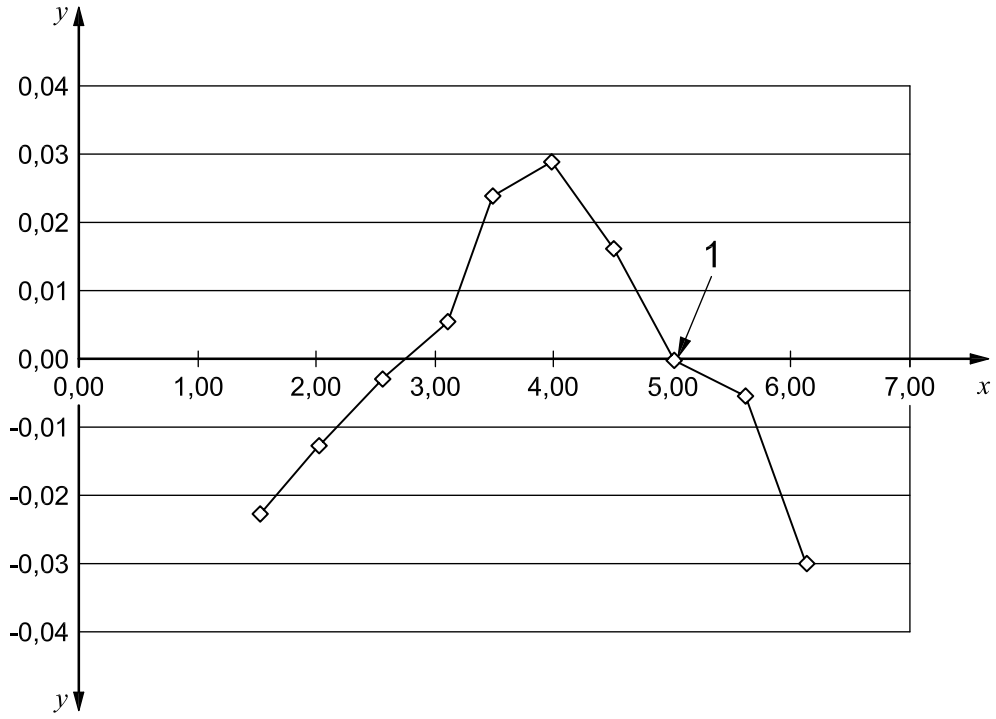


Key

1 instrumental results
2 linear regression line

x dilution, % mass per volume
y instrumental response

Figure C.1 — Linearity assessment: instrumental response against dilution



Key

1 linear regression residuals

x dilution, % mass per volume

y mean residuals

Figure C.2 — Linearity assessment: mean residuals against dilution

C.1.3.2 As a second example, linearity and upper limit data from a somatic cell counter (cf. ISO 13366-1|IDF 148-1[3]) are listed in Table C.4.

Table C.4 — Linearity and upper limit — Test sample set with progressive dilution of a high cell content milk by a low cell content milk)

Level N°	% dilution mass per volume	Measured concentration	Residuals reg. 1 to 21	Residuals reg. 1 to 9	Linearity ratio from 1 to i	Standard deviation	Student t-test for e_i vs 0	
i	x_i	y_i	e_i	e_i	$\Delta e/\Delta L$	$s_{yx,i}$	t_{obs}	
1	0,0	7,2	–25,2	–4,9		5,538	–1,01	
2	5,4	131,2	–18,2	–2,2	0,022	5,255	–0,46	
3	10,1	238,8	–12,4	–0,2	0,021	5,054	–0,04	
4	15,2	356,5	–5,1	3,0	0,023	4,884	0,65	
5	19,7	461,7	2,6	7,1	0,026	4,776	1,56	
6	24,4	564,0	3,1	3,8	0,022	4,704	0,85	
7	30,2	689,7	3,2	–0,7	0,018	4,675	–0,16	
8	35,0	800,2	9,7	2,0	0,015	4,699	0,43	
9	39,9	900,5	3,9	–7,8	0,017	4,770	–1,71	
10	44,9	1 013,5	8,6	–7,1	0,015	4,889	–1,54	
11	49,6	1 122,8	16,1	–3,4	0,013	5,045	–0,72	
12	55,3	1 249,3	19,1	–4,9	0,012	5,292	–1,02	
13	59,8	1 348,5	20,8	–6,8	0,011	5,530	–1,38	
14	64,5	1 441,7	12,2	–19,1	0,018	5,821	–3,80	← **upper limit** $t_{0,975} = 2,365$ with $p = 5$ % and $\nu = 7$
15	69,9	1 561,0	14,6	–21,1	0,018	6,208	–4,07	
16	74,6	1 653,5	5,3	–34,2	0,025	6,590	–6,39	
17	79,4	1 766,5	14,3	–29,0	0,023	7,024	–5,25	
18	84,6	1 865,2	0,4	–47,1	0,029	7,545	–8,23	
19	89,7	1 983,8	8,5	–43,0	0,027	8,105	–7,25	
20	95,5	2 074,8	–26,1	–82,3	0,043	8,804	–13,31	
21	100,0	2 143,0	–55,4	–115,2	0,057	9,391	–18,04	
Number, N	21	21	21	9		9	9	
Mean	49,89	1 113,02	0,00	0,00				
Standard deviation	30,95	670,56	18,957	4,905				
Minimum	0,00	7,20	–55,39	–7,80	0,02	4,67	–1,71	
Maximum	100,00	2 143,00	20,84	7,10	0,03	5,54	1,56	
Δ = max. – min.	100,00	2 135,80	76,23	14,90	0,01	0,86	3,27	

From the data of Table C.4:

a) the standard deviation of residual means: $s_e = 19,0$ (measured);

b) the standard deviation of repeatability: $s_r = 16,4$ (measured);

c) the standard deviation of level biases: $s_L = 16,423$ [from $s_L = (s_e^2 - s_r^2/n)^{1/2}$].

The $\Delta e/\Delta L$ ratio linearity test gave, over the whole range, i.e. $i = 1 \ldots 21$, a value of 76,23/2 135,8 = 0,036 > 0,02, from which it is concluded that linearity is inadequate.

The $\Delta e/\Delta L$ ratio linearity test gave, from $i = 1 \ldots 13$, a value of $0{,}011 < 0{,}02$, from which it is concluded that this part of the curve is linear.

This test is the simplest to apply, and is generally recommended for quick routine checks.

The bias from linearity test, $F_{\text{obs}} = \left(s_r^2 + ns_L^2\right)/s_r^2 = ns_e^2/s_r^2$, should be lower than $F_{0,95}$ with $k_1 = q - 2$ and $k_2 = q(n-1)$. With triplicate determinations on 21 levels, $k_1 = 19$, and $k_2 = 42$ gives $F_{\text{obs}} = 4{,}01 > F_{(0,95;19;42)} = 1{,}84$, from which it is concluded that linearity is inadequate.

The curve trend is assessed by comparison with second and third degree polynomial, using

$$F_{\text{obs}} = \left[(q-2)s_{yx}^2 - (q-k-1)s_{yx}k^2\right] / \left[(k-1)s_{yx}k^2\right] < F_{1-\alpha}$$

or equivalently:

$$s_{yx}/s_{yx}k < \left\{\left[F_{1-\alpha}(k-1) + (q-k-1)\right]/(q-2)\right\}^{1/2}$$

with q samples, k polynomial degrees, $k_1 = k - 1$, $k_2 = q - k - 1$ and α risk of error.

From the example, using linear reference values, y, calculated from mass per volume dilution ratios and the raw data of the linear part (sample 1 to 9), the polynomial equations are

Degree k	$y = b_3 x^3$	$+ b_2 x^2$	$+ b_1 x$	$+a$	s_{yx}^k
3	$y_3 = 2{,}45 \times 10^{-8} x^3$	$- 3{,}69 \times 10^{-5} x^2$	$+1{,}008\,9\,x$	+2,9	9,28
2	$y_2 =$	$4{,}15 \times 10^{-5} x^2$	$+0{,}945\,3\,x$	+11,4	11,05
1	$y_1 =$		$+1{,}036\,2\,x$	+32,390 894	19,65

and F-tests from the residual standard deviations s_{yx}^k and equivalent tests through residual standard deviations ratios are

Degree k	k_1	k_2	F_{obs}	$F_{(0,95;k_1;k_2)}$	**or**	s_{yx}/s_{yx}^k	**Limit**
3	2	17	34,1 (s[a])	> 3,59		2,12 (s[a])	> 1,15
2	1	18	42,1 (s[a])	> 4,41		1,78 (s[a])	> 1,09

[a] s = significant.

with, for $F_{(0,95;k_1;k_2)}$ values, $F_{(0,95;2;17)} = 3{,}59$ and $F_{(0,95;1;18)} = 4{,}41$.

Table C.5 — Assessment of a suitable equation for linearity adjustment

Level N°	Reference from % dilution	Mean concention	Predictions through equations			Residuals		
	y	x	$y_1(x)$	$y_2(x)$	$y_3(x)$	$y_1(x) - y$	$y_2(x) - y$	$y_3(x) - y$
1	12,1	7,2	0,8	18,2	10,2	18,6	1,1	9,2
2	133,4	131,2	118,9	136,2	134,7	14,5	–2,7	–1,3
3	239,0	238,8	228,2	239,5	242,1	10,8	–0,5	–3,1
4	353,5	356,5	346,7	353,7	359,0	6,8	–0,2	–5,5
5	454,6	461,7	451,4	456,7	463,3	3,2	–2,1	–8,7
6	560,2	564,0	560,7	557,8	564,6	–0,5	2,4	–4,5
7	690,4	689,7	695,5	683,1	689,3	–5,1	7,3	1,1
8	798,2	800,2	807,2	794,4	799,2	–8,9	3,8	–1,0
9	908,3	900,5	921,1	896,3	899,5	–12,8	12,0	8,8
10	1 020,6	1 013,5	1 037,4	1 012,1	1 013,2	–16,8	8,5	7,4
11	1 126,2	1 122,8	1 146,6	1 125,1	1 124,0	–20,5	1,1	2,1
12	1 254,2	1 249,3	1 279,2	1 257,1	1 253,7	–25,0	–2,9	0,5
13	1 355,3	1 348,5	1 383,8	1 361,6	1 356,6	–28,6	–6,3	–1,4
14	1 460,8	1 441,7	1 493,1	1 460,5	1 454,5	–32,3	0,4	6,4
15	1 582,1	1 561,0	1 618,7	1 588,1	1 581,4	–36,6	–6,0	0,7
16	1 687,7	1 653,5	1 728,0	1 687,8	1 681,4	–40,3	–0,2	6,3
17	1 795,5	1 766,5	1 839,6	1 810,7	1 805,5	–44,1	–15,2	–10,0
18	1 912,3	1 865,2	1 960,5	1 918,9	1 915,9	–48,2	–6,6	–3,6
19	2 026,8	1 983,8	2 079,1	2 049,9	2 051,1	–52,2	–23,1	–24,3
20	2 157,1	2 074,8	2 213,9	2 151,2	2 156,9	–56,8	5,9	0,2
21	2 258,2	2 143,0	2 318,6	2 227,6	2 237,4	–60,4	30,5	20,7
Number of data, N	21	21	21	21	21	21	21	21
Mean	1 132,69	1 113,02	1 153,75	1 132,69	1 133,03	–20,72	0,34	0,00
Standard deviation	695,07	670,56	718,99	694,99	694,44	24,50	10,39	8,87
Minimum	12,13	7,20	0,76	18,24	10,18	–60,41	–23,07	–24,26
Maximum	2 258,16	2 143,00	2 318,57	2 227,63	2 237,44	18,57	30,53	20,72
Δ = max. – min.	2 246,03	2 135,80	2 317,81	2 209,39	2 227,27	78,98	53,61	44,98
$\Delta e/\Delta L$						0,035	0,024	0,019

In conclusion, significant improvement is obtained by both second and third degree polynomials, which confirms linearity is inadequate. However, by applying the $\Delta e/\Delta L$ test on polynomial residuals with ratio values of 0,024 and 0,019 for second and third degree polynomials respectively, only the latter allows suitable adjustment with $\Delta e/\Delta L$ below 0,02.

C.1.4 Assessment of measurement limits

C.1.4.1 Assessment of upper limit

This subclause discusses the example of a somatic cell counter (cf. ISO 13366-1|IDF 148-1[3]).

On a regression residual plot vs concentration, calculation of the regression equation $y = bx + a$ on the linear part (levels 1 to 9), gives slope b = 22,460 3 and intercept a = 12,132 4.

Calculation of the residuals over the whole range and application of the Student t-test on the residuals provides a residual of 19,1 × 1 000 cells/ml with $t_{obs} = 3{,}80$ significant for $p = 0{,}95$ with seven degrees of freedom, which is higher than a Student t-value of 2,87. Level 14 corresponds to the upper limit.

It can be concluded that deviation from linearity occurs from level 14, whose concentration constitutes the upper limit of measurement.

C.1.4.2 Lower limit

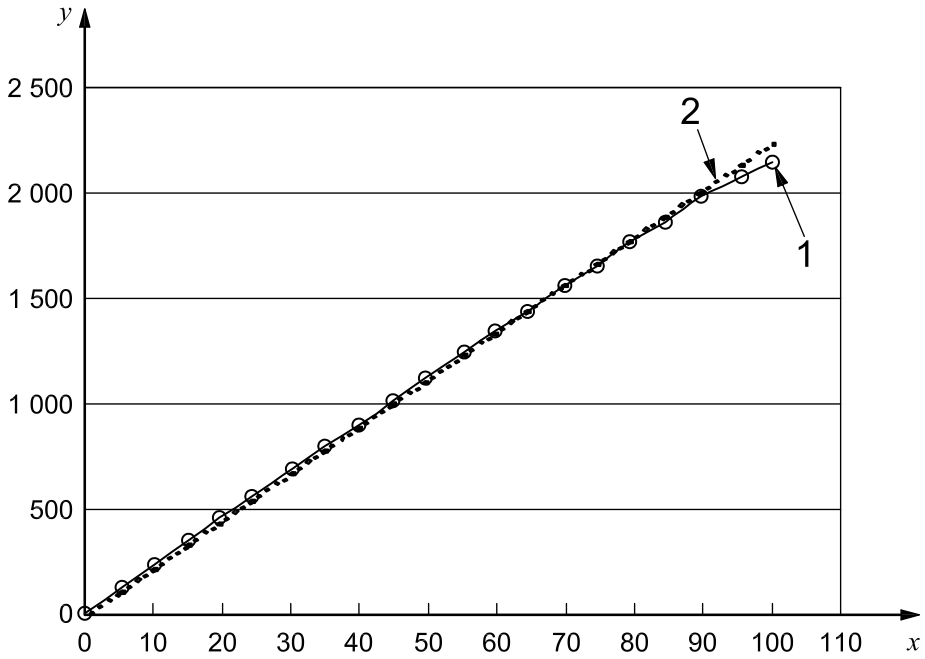
Assessment of the determination limit, L_{det}, or the quantification limit, L_Q, is done on 10 measurements of a set of somatic cell counting data close to zero, using limits of Table B.1, according to the following example.

Data: 3 - 5 - 4 - 3 - 5 - 4 - 5 - 3 - 5 - 4

Mean, $\bar{x}$	=	4,100	
Std deviation, σ_x	=	0,876	
CV, %	=	21,4	$< 30\ \% \Rightarrow$ conformity
$L_{det,\alpha,\beta} = 5\ \%$	=	2,881	$< 5{,}000 \Rightarrow$ conformity
$L_{det,\alpha,\beta} = 1\ \%$	=	4,513	$< 5{,}000 \Rightarrow$ conformity
$L_{det,\alpha,\beta} = 0{,}13\ \%$	=	5,256	$> 5{,}000 \Rightarrow$ non-conformity
Number, N	=	10	

It can be concluded that the determination limit complies with the limit 5 000 cells/ml when using accepted risks of error α and β of minimum 1%.

See Figures C.3 and C.4.

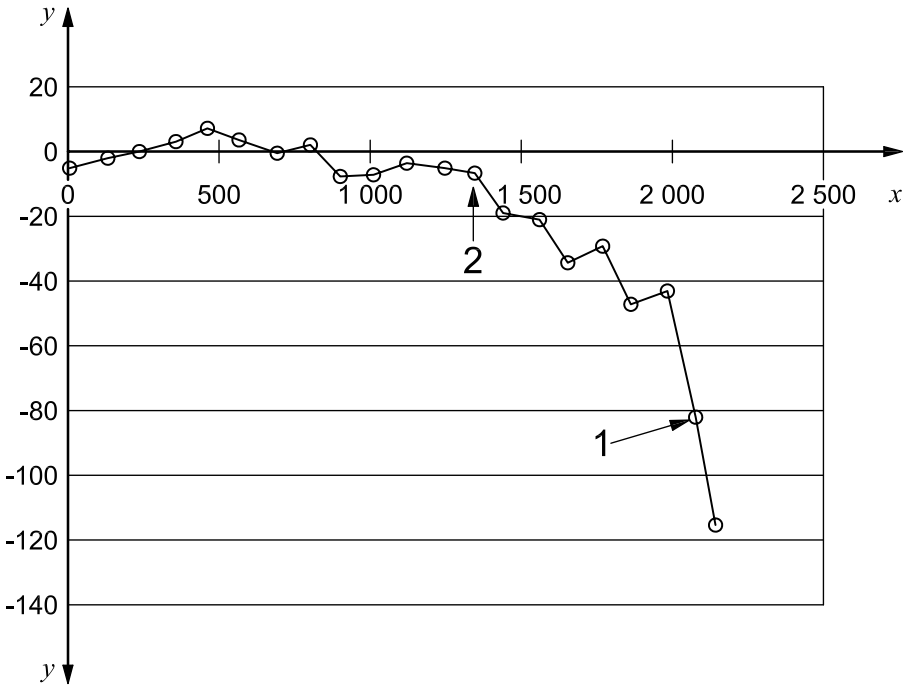


Key

1 measured concentration
2 theoretical concentration

x dilution, % mass per volume
y concentration, 1 000 cells/ml

Figure C.3 — Linearity curve: cell concentration against dilution



Key

1 bias from linearity (regression from 0 to 900)

2 upper limit

x mean concentrations, 1 000 cells/ml

y residuals, 1 000 cells/ml

Fitting of the curve in the linear part: 0 to 900 000.

Figure C.4 — Deviation from linearity: distribution of residuals throughout the concentration range

C.2 Assessment of the overall accuracy

EXAMPLE Fat analysed by infra red spectroscopy (cf. ISO 9622) using a set of individual milk samples.

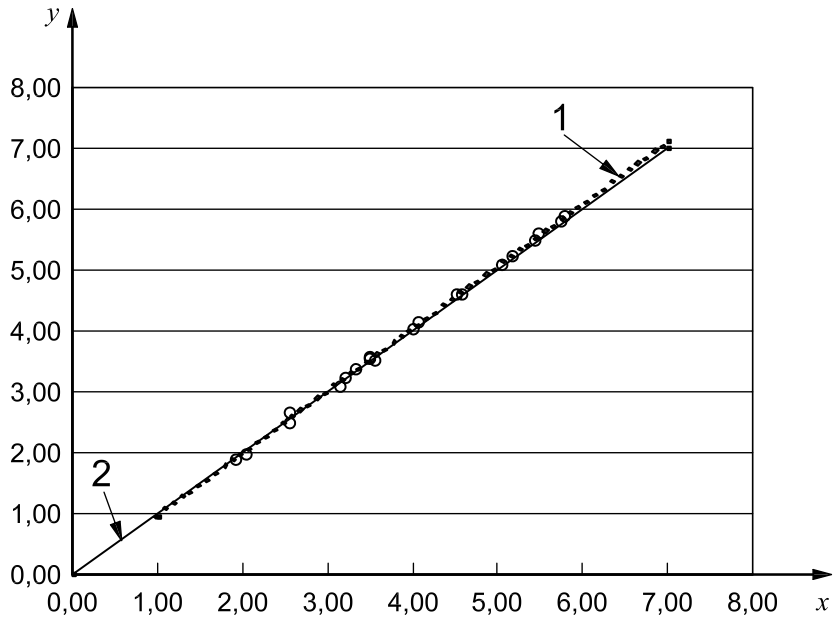
Table C.6 — Example of fat analyses

Sample N°	Reference method	Instrumental method				Repeatability	Accuracy	
		Test 1	Test 2	Mean	Corrected	Range	Bias	Residual
	y	x_1	x_2	$\bar{x}$	$y(x)$	$w = \lvert x_1 - x_2 \rvert$	$d = x - y$	$e = y - y(x)$
1	1,89	1,92	1,94	1,930	1,90	0,02	0,04	–0,006
2	1,98	2,05	2,06	2,055	2,03	0,01	0,07	–0,045
3	2,48	2,55	2,56	2,555	2,54	0,01	0,07	–0,061
4	2,66	2,56	2,56	2,560	2,55	0,00	–0,10	0,114
5	3,10	3,16	3,13	3,145	3,15	0,03	0,04	–0,049
6	3,23	3,20	3,22	3,210	3,22	0,02	–0,02	0,014
7	3,37	3,31	3,34	3,325	3,33	0,03	–0,04	0,035
8	3,57	3,51	3,50	3,505	3,52	0,01	–0,06	0,050
9	3,53	3,51	3,50	3,505	3,52	0,01	–0,02	0,010
10	3,52	3,57	3,57	3,570	3,59	0,00	0,05	–0,067
11	4,02	4,00	4,01	4,005	4,04	0,01	–0,01	–0,016
12	4,15	4,05	4,09	4,070	4,10	0,04	–0,08	0,047
13	4,59	4,52	4,51	4,515	4,56	0,01	–0,08	0,028
14	4,61	4,59	4,57	4,580	4,63	0,02	–0,03	–0,019
15	5,10	5,06	5,06	5,060	5,12	0,00	–0,04	–0,024
16	5,23	5,18	5,19	5,185	5,25	0,01	–0,04	–0,022
17	5,49	5,44	5,44	5,440	5,52	0,00	–0,05	–0,025
18	5,61	5,48	5,47	5,475	5,55	0,01	–0,14	0,058
19	5,80	5,74	5,76	5,750	5,84	0,02	–0,05	–0,035
20	5,89	5,80	5,78	5,790	5,88	0,02	–0,10	0,014
Number of data, N	20	20	20	20	20	20	20	20
Mean	3,991	3,960	3,963	3,962	3,991	0,014	–0,030	0,000
Standard deviation	1,260	1,223	1,219	1,221	1,259	0,011	0,059	0,047
Minimum	1,890	1,920	1,940	1,930	1,896	0,00	–0,14	–0,07
Maximum	5,890	5,800	5,780	5,790	5,876	0,04	0,07	0,11
Δ = max. – min.	4,000	3,880	3,840	3,860	3,980	0,04	0,21	0,18

In conclusion, instrument accuracy complies with limits defined for the component analysed and the type of milk fat in individual cow milk, by reference to the protocol and Clause B.1, thus is assessed positively. The mean bias and slope significantly differ from 0 and 1 and indicate that calibration can still be optimized further if needed for the purpose.

See Figures C.5 and C.6.

	Parameter	Estimate	Limits	Conformity
Repeatability	s_r	0,012	0,014	Yes
Accuracy	**Mean bias**, $\overline{d}$	–0,030	± 0,050	Yes
	s_d (= s_{x-y})	0,059	0,100	Yes
	N	20		
	t_{obs} **versus 0**	2,218	$t_{0,975} = 2,093$	$p > 0,05$
	ν	19		
Regression	**Slope**, b	1,031 1	1 ± 0,05	Yes
	s_b	0,008 8		
	t_{obs} b **versus 1**	3,511	$t_{0,975} = 2,101$	$p > 0,001$
	Intercept, a	–0,093 5		
	s_a	0,037		
	t_{obs} a **versus 0**	2,556	$t_{0,975} = 2,101$	$p > 0,05$
	ν	18		
	s_{yx}	0,047	0,100	Yes



Key

1 regression line

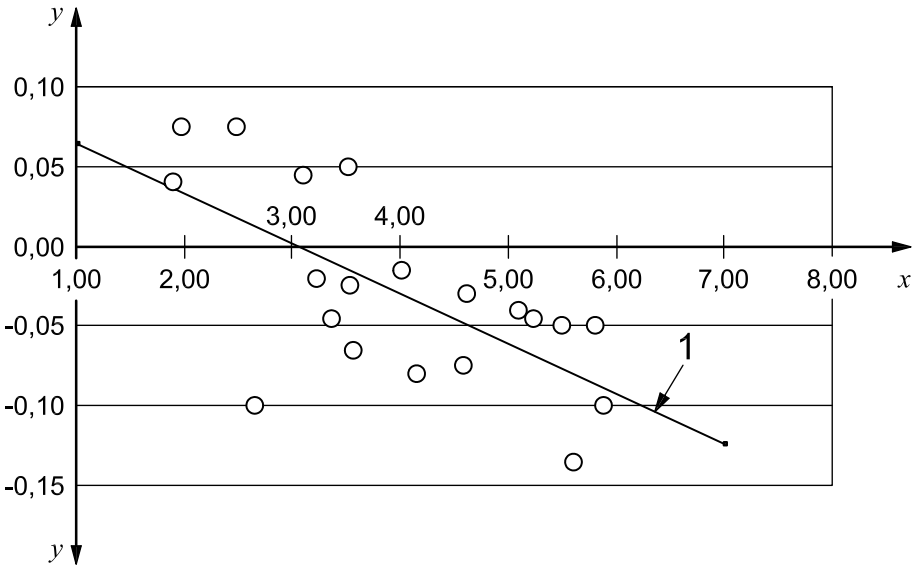
2 theoretical line, $y = x$

x instrumental method, fat content, % mass fraction

y reference method, fat content, % mass fraction

○ mean x

Figure C.5 — Assessment of accuracy: results for the reference method plotted against those for the instrumental method



Key

1 regression line

x reference method, fat content, % mass fraction

y difference between instrument and reference, fat content, % mass fraction

○ differences, $d = x - y$

$y = 0{,}031\ 4x + 0{,}095\ 8$

$s_{x-y} = 0{,}059$

$s_{yx} = 0{,}047$

Figure C.6 — Assessment of accuracy: distribution of differences to the reference method

Annex D
(informative)

Summary of statistical formulas for method evaluations

D.1 Assessment of the precision

Standard deviation of repeatability: (q levels and n replicates)

$$s_r = \left\{ \frac{\sum_{i=1}^{q} \left[\sum_{j=1}^{n} \left(x_{ij} - \bar{x}_i \right)^2 \right]}{q(n-1)} \right\}^{1/2}$$

and with $n = 2$ $\quad s_r = \left[\sum_{i=1}^{q} \frac{\left(x_{i1} - x_{i2} \right)^2}{2q} \right]^{1/2}$

or $\quad s_r = \left[\sum_{i=1}^{q} \frac{w_i^2}{2q} \right]^{1/2}$

Standard deviation of means of control test checks:

$$s_{\bar{x}}^2 = \frac{\left[\sum x_j^2 - \left(\sum \frac{\bar{x}_j^2}{q} \right) \right]}{q-1}$$

Standard deviation between control test checks:

$$s_c = \left(\frac{s_{\bar{x}}^2 - s_r^2}{n} \right)^{1/2}$$

Standard deviation of daily reproducibility:

$$s_R^2 = s_c^2 + s_r^2 = s_{\bar{x}}^2 + s_r^2 \left(1 - \frac{1}{n} \right)$$

D.2 Assessment of the accuracy

Means:

$$\bar{x} = \frac{\sum x_i}{q}$$

$$\bar{y} = \frac{\sum y_i}{q}$$

$$\bar{d} = \frac{\sum d_i}{q} = \frac{\sum x_i - \sum y_i}{q} = \bar{x} - \bar{y}$$

Sum of squares and of products:

$$S_x = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \left(\frac{\sum x_i^2}{q} \right)$$

$$S_y = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \left(\frac{\sum y_i^2}{q} \right)$$

$$S_d = \sum (d_i - \bar{d})^2 = \sum d_i^2 - \left(\frac{\sum d_i^2}{q} \right)$$

$$P_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \left(\frac{\sum x_i \sum y_i}{q} \right)$$

Slope:

$$b = \frac{P_{xy}}{S_x}$$

Intercept:

$$a = \bar{y} - b\bar{x}$$

Estimate for x:

$$\bar{y}(x) = b\bar{x} + a$$

Conditional mean for x:

$$y(x) = bx + a$$

Residual, e_i:

$$e_i = y_i - \bar{y}(x_i) = y_i - bx_i - a$$

Difference, d_i:

$$d_i = x_i - y_i$$

Correlation coefficient:

$$r_{xy} = \left(\frac{P_{xy}^2}{S_y S_x} \right)^{1/2}$$

Standard deviations of

— **differences, d:**

$$s_d = \left(\frac{S_d}{q-1} \right)^{1/2}$$

$$s_d = \left(\frac{S_y^2 + S_x^2 - 2P_{xy}}{q-1} \right)^{1/2}$$

— **residuals, e_i:**

$$s_{yx} = \left(\frac{\sum (y_i - bx_i - a)}{q-2} \right)^{1/2}$$

$$s_{yx} = \left[\frac{s_y^2 - (P_{xy}^2 / s_x)}{q-2} \right]^{1/2}$$

$$s_{yx} = \left[\frac{s_y^2 (1 - r_{xy}^2)}{q-2} \right]^{1/2}$$

— **slope, b:**

$$s_b = \left(\frac{s_{yx}}{S_x} \right)^{1/2} = \frac{s_{yx}}{s_x (q-1)^{1/2}}$$

— **intercept, a:**

$$s_a = s_{yx} \left(\frac{1}{q} + \frac{\bar{x}^2}{S_x} \right)^{1/2} = s_{yx} \left[\frac{1}{q} + \frac{\bar{x}^2}{(q-1)s_x^2} \right]^{1/2}$$

— **conditional mean, $\bar{y}(x_0)$:**

$$s_{\bar{y}(x_0)} = s_{yx} \left[\frac{1}{q} + \frac{(x_0 - \bar{x})^2}{S_x} \right]^{1/2} = s_{yx} \left[\frac{1}{q} + \frac{(x_0 - \bar{x})^2}{(q-1)s_x^2} \right]^{1/2}$$

— **single estimate, $y(x_0)$:**

$$s_{y(x_0)} = s_{yx} \left[1 + \frac{1}{q} + \frac{(x_0 - \bar{x})^2}{S_x} \right]^{1/2}$$

D.3 Conformity tests

D.3.1 Conformity of an estimate

a) slope b versus 1,000:

$$t_{\text{obs}} = \frac{|b - 1{,}000|}{s_b} \leqslant t_{1-\alpha/2}$$

with $\nu = q - 2$ and $\alpha = 0{,}05$

b) slope b versus 0,00:

$$t_{\text{obs}} = \frac{|b|}{s_b} \leqslant t_{1-\alpha/2}$$

with $\nu = q - 2$ and $\alpha = 0{,}05$

c) mean difference $\overline{d}$ versus 0,00:

$$t_{\text{obs}} = \frac{|\overline{d}|\sqrt{q}}{s_d} \leqslant t_{1-\alpha/2}$$

with $\nu = q - 1$ and $\alpha = 0{,}05$

or $\overline{x}$ versus $\overline{y}$ (when $b \neq 1{,}000$):

$$t_{\text{obs}} = \frac{|\overline{x} - \overline{y}|\sqrt{q}}{s_{yx}} \leqslant t_{1-\alpha/2}$$

with $\nu = q - 2$ and $\alpha = 0{,}05$

d) intercept a versus 0,00:

$$t_{\text{obs}} = \frac{|a|}{s_a} \leqslant t_{1-\alpha/2}$$

with $\nu = q - 2$ and $\alpha = 0{,}05$

e) conditional mean, $\overline{y}_{x_0}$, versus reference value y_0, or residual e_0, versus 0,00:

$$e_0 = y_0 - \overline{y}_{x_0} = y_0 - b_{q-1}\, x_0 - a_{q-1} \text{ and}$$

$$s_{y(\overline{x}_0)} = s_{yx,q-1}\left[\frac{1}{q-1} + \frac{(x_0 - \overline{x}_{q-1})^2}{S_{x_{q-1}}}\right]^{1/2}$$

$$t_{\text{obs}} = \frac{|e_0|}{s_{y(\overline{x}_0)}} \leqslant t_{1-\alpha/2}$$

with $\nu = q - 3$ and $\alpha = 0{,}05$

For outlier detection or departure from linearity, check whether point M_o (x_o,y_o) belongs to the linear curve calculated without that point.

D.3.2 Conformity of a standard deviation, s, versus an upper limit, σ

a) method 1, χ^2:

$$\sigma^2 \geqslant \frac{k s^2}{\chi^2_\alpha} \quad \Rightarrow \quad s \leqslant \sigma\left(\frac{\chi^2_\alpha}{k}\right)^{1/2}$$

with $\nu = k$ and $\alpha = 0{,}05$

b) method 2 (error standard) which can replace method 1 for $k > 50$)

$$s + \frac{u_{1-\alpha} s}{\sqrt{2k'}} \leqslant \sigma \quad \Rightarrow \quad s \leqslant \frac{\sigma}{(1 + u_{1-\alpha}/\sqrt{2k'})}$$

with k' data and $\alpha = 0{,}05$

D.3.3 F-test of one-way ANOVA for factor f measured by $\overline{x}_i$

$$F_{\text{obs}} = \frac{\left(n s_f^2 + s_r^2\right)}{s_r^2} = \frac{n\left(s_{\overline{x}}^2 - s_r^2/n\right) + s_r^2}{s_r^2} = \frac{n s_{\overline{x}}^2}{s_r^2} < F_{1-\alpha}$$

or

$$\frac{s_{\overline{x}}}{s_r} < \left(\frac{F_{1-\alpha}}{n}\right)^{1/2}$$

where

n is the number of replicates;

$\bar{x}_i$ are the means of replicates;

$k_1 = q - 1$

$k_2 = q(n - 1)$

α is the risk of error.

NOTE Factor f, related to s_f, can be sample, $s_f = s_s$, for batch homogeneity assessment or checks, $s_f = s_c$, or for method stability.

D.3.4 Linearity tests (significant lack-of-fit)

D.3.4.1 Sample or level (s_l) effect compared to repeatability (s_r) interpreted as linearity defect:

$$F_{\text{obs}} = \frac{(ns_L^2 + s_r^2)}{s_r^2} = \frac{\left[n\left(s_{\bar{e}}^2 - s_r^2/n\right) + s_r^2\right]}{s_r^2} = \frac{ns_{\bar{e}}^2}{s_r^2} < F_{1-\alpha}$$

or

$$\frac{s_{\bar{e}}}{s_r} < \left(\frac{F_{1-\alpha}}{n}\right)^{1/2}$$

where

$\bar{e}$ is the mean residual of n replicates;

$k_1 = q - 2$;

$k_2 = q(n - 1)$;

α is the risk of error.

D.3.4.2 Curve trend through comparing regression residual errors of a line with a k degree polynomial:

$$F_{\text{obs}} = \frac{(q-2)s_{yx}^2 - (q-k-1)s_{yx}k^2}{(k-1)s_{yx}k^2} < F_{1-\alpha}$$

or

$$\frac{s_{yx}}{s_{yx}k} < \left[\frac{F_{1-\alpha}(k-1) + (q-k-1)}{q-2}\right]^{1/2}$$

with: q samples, k polynomial degrees, $k_1 = k - 1$, $k_2 = q - k - 1$ and α risk of error.

Annex E
(informative)

Procedure for sample set preparation in linearity evaluation

E.1 General

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

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.

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

E.2 Reference for linearity

Reference values for linearity samples can be established from either the mixing ratio or the theoretical concentrations as calculated from the concentrations of the initial samples.

Depending on the alternative method, they should be obtained from volume fraction mixing ratios, where analysis is performed on a milk volume (volumetric intake measurement), and mass fraction mixing ratios, where analysis is applied to a weighed milk portion.

Mixing liquids of different densities can lead to erroneous conclusions due to a non-linear relationship between (mass fraction) and (volume fraction) ratios. Both ratios can be used equivalently, as equal, only if the respective densities of both liquids are identical.

When using weighing for more accurate quantity measurements, the mass fraction, w, can be transformed into the appropriate volume fraction, φ, according to the equation:

$$\varphi = \frac{w}{w(1 - d_H / d_L) + d_H / d_L}$$

where

d_H is the mass of the volume of sample L_H;

d_L is the mass of the volume of sample L_L.

E.3 Sample set preparation

E.3.1 General

Select or prepare two laboratory samples of liquid materials with respectively high and low component contents so as to cover the entire range of assessment. Determine the concentration of the component in the stock liquid materials, L_H and L_L, and the range of concentrations $\Delta L_{test} = L_H - L_L$.

Define the suitable level (sample) number, n, and determine the constant concentration increment,

$$\Delta L_{\text{increment}} = \frac{L_{\text{H}} - L_{\text{L}}}{n-1}$$

needed to arrive at an even spread of the n samples across the range.

For sample $i = 1 \ldots n$, calculate the theoretical ratios,

$$R_i = \frac{i-1}{n-1}$$

and the related target concentrations of the linearity sample set,

$$F_i = L\,\frac{(i-1)\,(L_{\text{H}} - L_{\text{L}})}{n-1}$$

to check the appropriateness of tentative theoretical reference values.

Proceed to subsequent quantity measurements by weighing and mixing using the theoretical ratios. Then, from the quantities measured, calculate the actual mixing ratios and the theoretical reference values as in E.3.2 and E.3.3.

E.3.2 Volumetric milk measurement

Volume ratios: $R_{V_i} = V_{\text{H}_i} / V_{F_i}$

From weighing: $R_{V_i} = (m_{\text{H}} / d_{\text{H}}) / [(m_{\text{H}} / d_{\text{H}}) + (m_{\text{L}} / d_{\text{L}})]$ or

$$R_{V_i} = R_m / \left\{ R_{m_i} \left[1 - (d_{\text{H}} / d_{\text{L}}) \right] + (d_{\text{H}} / d_{\text{L}}) \right\}$$

Particular case: $R_{V_i} = R_{m_i}$ when $m_{\text{H}} = m_{\text{L}}$

Theoretical final concentration: $F_i = R_{V_i}(L_{\text{H}} - L_{\text{L}}) + L_{\text{L}}$, with L_{H}, L_{L} and F_i expressed in mass or volume units.

E.3.3 Weighed milk measurement

Mass ratios: $R_{m_i} = m_{\text{H}_i} / m_{F_i}$

Theoretical final concentration: $F_i = (m_{\text{H}_i} / m_{F_i})\,(L_{\text{H}} - L_{\text{L}}) + L_{\text{L}}$ with L_{H}, L_{L} and F_i expressed as mass fractions

where

d_{H} and d_{L} are the relative densities of samples L_{H} and L_{L} respectively, and

V_{H_i}, V_{F_i}, m_{H_i}, m_{F_i} are the volumes and masses measured for the high content sample and the final mixtures of the ith level.

EXAMPLE

Starting from:

1) a high content liquid material of component concentration with $L_H = 80$ units and a relative density of $d_H = 1{,}06$ (e.g. skim milk retentate for protein) at 20 °C;

2) a low content liquid material of component concentration with $L_L = 10$ units and a relative density of $d_L = 1{,}02$ (e.g. milk ultrafiltrate) at 20 °C; and

3) a sample number of $n = 10$ levels.

The concentration range in the sample set is then $\Delta L_{test} = L_H - L_L = 70$ units and the concentration increment 7,8 units.

For a given final quantity (either volume or mass), respective quantities of high, Q_{H_i}, and low, Q_{L_i}, to be mixed for an even spread of the concentrations F_i are listed in Table E.1.

Table E.1 — Theoretical and targeted quantities

Preparation Level	Targeted/Theoretical			
	R_i [a]	F_i	Q_{H_i} [a]	Q_{L_i} [a]
$i = 1$ to n	$(i-1)/(n-1)$	$L_L + (i-1)(L_H - L_L)/(n-1)$	$R_i\ Qf$	$Qf - Q_{H_i}$
1	0,000	10,0	0,0	100,0
2	0,111	17,8	11,1	88,9
3	0,222	25,6	22,2	77,8
4	0,333	33,3	33,3	66,7
5	0,444	41,1	44,4	55,6
6	0,556	48,9	55,6	44,4
7	0,667	56,7	66,7	33,3
8	0,778	64,4	77,8	22,2
9	0,889	72,2	88,9	11,1
10	1,000	80,0	100,0	0,0

[a] R and Q are measured on either a volume or a mass basis.

Measured quantities as listed in Table E.1 are mixed at a temperature at which the assumed mass of volumes is valid (about 20 °C to 25 °C). Due to approximated quantities and possible small measurement errors, the true mixing ratios and relevant theoretical concentration are recalculated.

If volumes are measured by weighing (for a better precision of the reference), further transformation using the mass of volumes is applied to calculate the volumes of test portion intake, volume ratios and theoretical values used as linearity reference.

If the test portion intake is weighed, mass values serve to recalculate the mass fraction ratios and related theoretical calculated values as in Table E.2.

Table E.2 — Quantity ratios and theoretical reference

Actual Level	Masses measured		Volumes calculated		Quantity ratios		Theoretical reference[a]	
	m_{H_i}	m_{L_i}	V_{H_i}	V_{L_i}	R_{m_i}	R_{V_i}	F_{m_i}	F_{V_i}
$i = 1$ to n			m_{H_i}/d_H	m_{L_i}/d_L	$m_H/(m_H + m_L)$	$V_H/(V_H + V_L)$	$R_{m_i}.(m_H - m_L) + m_L$	$R_{V_i}.(V_H - V_L) + V_L$
1	0,0	102,0	0,0	100,0	0,000	0,000	10,0	10,0
2	11,5	89,0	10,8	87,3	0,114	0,111	18,0	17,7
3	21,9	78,0	20,7	76,5	0,219	0,213	25,3	24,9
4	33,1	67,3	31,2	66,0	0,330	0,321	33,1	32,5
5	44,2	55,9	41,7	54,8	0,442	0,432	40,9	40,2
6	56,2	44,1	53,0	43,2	0,560	0,551	49,2	48,6
7	67,0	33,4	63,2	32,7	0,667	0,659	56,7	56,1
8	78,1	22,0	73,7	21,6	0,780	0,774	64,6	64,1
9	89,0	10,9	84,0	10,7	0,891	0,887	72,4	72,1
10	100,3	0,0	94,6	0,0	1,000	1,000	80,0	80,0

[a] Depending on the principle of milk test portion intake.

Bibliography

[1] ISO 1211|IDF 1, *Milk — Determination of fat content — Gravimetric method (Reference method)*

[2] ISO 8968|IDF 20 (all parts), *Milk — Determination of nitrogen content*

[3] ISO 13366-1|IDF 148-1, *Milk — Enumeration of somatic cells — Part 1: Microscopic method (Reference method)*

[4] ISO 13366-2, *Milk — Enumeration of somatic cells — Part 2: Guidance on the operation of fluoro-opto-electronic counters*

[5] ISO 14637|IDF 195, *Milk — Determination of urea content — Enzymatic method using difference in pH (Reference method)*

[6] ISO 16140, *Microbiology of food and animal feeding stuffs — Protocol for the validation of alternative methods*

[7] ISO 17997|IDF 29 (all parts), *Milk — Determination of casein-nitrogen content*

[8] ISO 22662|IDF 198, *Milk and milk products — Determination of lactose content by high-performance liquid chromatography (Reference method)*

[9] ISO/IEC Guide 99:2007, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*

ICS 67.100.10

Price based on 45 pages