IDENTIFICATION OF ACID NEUTRALIZERS IN FRAUDULENT MILK: FULL VALIDATION OF A CLASSICAL QUALITATIVE METHOD

Carina de Souza Gondim*, Marina Penna e Palhares*, Pedro Paulo Borges dos Santos*, Roberto Cesar de Sousa*, Roberto Gonçalves Junqueira* e Scheilla Vitorino Carvalho de Souza*,a,b,3

aDepartamento de Alimentos, Faculdade de Farmácia, Universidade Federal de Minas Gerais, Pampulha, 31270-010 Belo Horizonte – MG, Brasil

Received on 12/08/2020; accepted on 11/12/2020; published on web on 20/01/2021

Neutralization with alkaline compounds is one of the most common adulterations in milk. The rosolic acid method is a classical test widely used in different countries for detection of neutralizers in milk. The official and a modified version were validated in a single laboratory validation process considering four adulterants: sodium bicarbonate (BI), sodium carbonate (CA), sodium hydroxide (HY) and sodium citrate (CI). The modified version, which presented better performance was selected for interlaboratory validation. In this process, samples of raw milk with acidities of 0.19% were neutralized with different concentrations of BI, CA, HY and C and tested for homogeneity and stability. Eight laboratories, which represented different sectors of the milk production chain, received and analysed these samples. The collaborative trial results confirmed the method performance, although sensitivity and precision were inferior to those obtained in the intralaboratory process, demonstrating its applications and limitations.

Keywords: collaborative trial; milk adulteration; rosolic acid; single laboratory validation; screening method.

INTRODUCTION

Neutralization with alkaline compounds is a common adulteration made to mask the alteration of milk suffering from excess acidity, which may cause serious health-related problems. The increased acidity caused by bacterial growth and lactic acid production is often associated with failures in good production practices that may occur along the production chain, mainly from the milking to receiving in the dairy industry steps.1-3

The equilibrium of ionized and ionizable components has significant importance in dairy chemistry, in which different substances such as carbon dioxide, proteins, citrates, lactates and phosphates act. These substances, principally at pH values between 5 and 6, act as buffers,4 that is, they can withstand variations in pH when there is the addition of acid or base or with the dilution of the solution. The addition of neutralizing substances, however, can alter the chemical composition of milk as well as modify its quality, and it can consequently create economic problems for the dairy chain in addition to posing risks to the health of consumers.5,6

The acidity neutralizer tests compose the list of analyses that must be carried out daily by the industries to control the quality of the raw milk.7,8 The rosolic acid and phenolphthalein qualitative methods have been described for this purpose in the legislation of several countries.8,9

The rosolic acid method is based on acid-base reactions. Therefore, the addition of neutralizing compounds could not be identified by this test if the amount of adulterant added to the milk result in a sample with a pH value and acidity within the allowed limits.10 Different pH ranges for chemical species of this substance have been reported in the literature.11 According to Ritter,11 rosolic acid has yellow staining at acidic pHand red color at basic pH with a turning point close to a pH of 7. Tronco10 argued that rosolic acid turns at pH values between 6.8 and 8.9.

In common with the other qualitative tests for milk quality control, these methods are widely used in routine analyses in milk quality controls in several countries and studies involving their performance have received little attention.8,9,12 In general, the presence of these compounds in milk is prohibited. However, these legislations do not have a limit for which such substances could be detected in raw milk. Also, the publications of the official methods do not give any information about the performance parameters, such as detection limits. In addition, the few works available in the literature do not constitute an appropriate validation process, especially if the current trends on the validation of qualitative methods are considered.13-17

The innovations of the present study are the proposition of a modified method of rosolic acid for the detection of acid neutralizers in raw milk aiming improve method performance and the strategies for a full validation processes of qualitative methods, involving single and interlaboratorial procedures.

MATERIALS AND METHODS

Reagents and equipment

Equipment and measuring instruments used were calibrated by accredited laboratories according to ISO/IEC 17025 and included the following: scales (AUX 220, Shimadzu; e BK 300, Gehaka), volumetric flasks, a thermostatic bath (314-8DN, Nova Ética), a cryoscope (PZL 900, PZL), a milk densimeter (5784, Incotherm), micropipettes (LM 10000, Soluções Lab HTL; Finnpipette F3, Thermo Scientific; Transferpette S, Bre) and thermometers (7665.02.0.00, Incotherm).

All compounds used were of analytical grade; in some instances, chemicals were purchased from multiple suppliers to study supplier influence on method of testing. Rosolic acid, ethyl alcohol, sodium carbonate (CA) and sodium hydroxide (HY) were supplied by Vetec Química Fina Ltda. (Duque de Caxias, RJ, Brazil), Dynamics Química Contemporânea Ltda. (Diadema, SP, Brazil) was the supplier of rosolic acid, sodium citrate (CI), phenolphthalein and HY. Sodium bicarbonate (BI) and phenolphthalein were purchased from Alphatec (São Bernardo do Campo, SP, Brazil). Ethyl alcohol was also obtained from FMAIA (Belo Horizonte, MG, Brazil).
Methods

In the rosolic acid method, the presence of neutralizers in milk samples is revealed by the action of the rosolic acid indicator. A test tube containing 5 mL of milk is added with 10 mL of an ethyl alcohol solution (98%). After manual shaking and adding of two drops of rosolic acid 2% solution in ethyl alcohol, the formed color is observed. The observation of red-carmine staining is the criterion of a positive result for the presence of neutralizing compounds (Figure 1).\(^9\)

An optimization step was performed to train analysts and to detect critical points in the analytical procedure that could influence the results and the classification of the samples.

Negative (blank) and positive (spiked) samples were analysed according to the positive criterion described in the legislation. Several analytical batches with different concentration ranges were investigated to identify the concentrations ranges that provided the observation of false-negative rates (FNRs) between 0 and 100%, indicating the adequacy of the concentration ranges for validation.\(^18\) Thus, the studied analytes were as follows: BI (between 0.0025 and 0.0500%, w/v), CA (0.0025-0.0500%, w/v), CI (0.03-0.15%, w/v) and HY (0.0025-0.0275%, w/v), with five concentrations levels and ten replicates per level (Figure 2a).

With the results obtained in this step, a new method based on the classical rosolic acid method was proposed. This modified method was the consequence of simple changes made in the analytical procedure of the classical method, including a new criterion for positive results, and was performed to help analysts in decision making and to allow an increased performance in detection of acid neutralizers (Figure 1).

In addition, the visualized colors for positive and negative samples were compared with colors presented by Krause,\(^19\) with the aim of establishing a color pattern in the red, green and blue (RGB) system for the results. The color scale obtained was used as a complementary tool by the analysts to classify the samples in both validation processes - single laboratory and collaborative trial. In this way, the collaborators could have a clearer idea of the colors named salmon or pink, for example.

Single laboratory validation

The official and modified methods were validated by a single laboratory according to the procedure proposed by Gondim et al.\(^18\)

Samples

At each experimental stage, raw milk samples (5 L) were collected from a refrigerated storage tank of the Pedro Leopoldo Farm of the Veterinary School of the Federal University of Minas Gerais (EV/UFMG). These samples were homogenized, packed in polypropylene gallons, identified, and refrigerated (2 to 7 °C) during transportation to the laboratory and up to the beginning of the experiments. The average time between sampling and the beginning of the experiments was 4 hours.

Samples were within the allowed limits for density and cryoscopy. The acidity of samples employed in the performance parameters estimation ranged from 0.17 to 0.19 mg L\(^{-1}\). The determination of

![Figure 1. Schematic representation of the rosolic acid qualitative methods for the detection of acid neutralizers in milk: official and modified methods (modified steps in gray)](image-url)
density, cryoscopy and acidity of samples was carried out in triplicate according to the official methods.\(^9\)

In density determination, a densimeter was immersed in a graduated cylinder containing 250 mL of milk sample. The value observed in the graduated scale, resulting from the displacement of the liquid, was taken as the density of the sample. The acidity determination consisted in the titration of a 20 mL of milk by a sodium hydroxide solution of known concentration, using phenolphthalein as indicator. The milk cryoscopy, corresponding to the measurement of its freezing point, was determined using an electronic cryoscope. In this equipment, the sample is rapidly cooled to a few degrees below its freezing point under constant stirring. The resulting vibration causes a thermal imbalance inside the sample, causing the releasing of the fusion heat. Then, the temperature increases until it reaches the freezing point, remaining constant for some time. During this period, the machine registers the freezing point.\(^9\)

The analyses of the samples occurred within the maximum interval of 48 hours after milking, the maximum allowed time for receiving raw milk by the industries as established in legislation,\(^7\) simulating the ideal conditions for milk being received by dairy industries.

**Preliminary tests**

The concentration range evaluated in the following steps of the validation process was defined after the analysis of positive samples that were spiked at five concentration levels of each analyte, with 10 replicates per level. This step was carried out in a single analytical batch using the same concentrations used in the optimization step (Figure 2a).\(^18\)

**Evaluation of rates, unreliability region, detection limit, accordance and concordance**

Samples (positive and negative) were prepared in 30 independent replicates. The concentrations of each analyte were selected based on the concentration ranges evaluated in the preliminary tests, using a higher number of concentration levels, equally spaced, for a better estimative of the analytical parameters, such as the URs. For BI and CA, the concentration range was set between 0.0025 and 0.0500% (w/v) levels, with 20 concentration levels. Eleven concentrations...
between 0.0025 and 0.0275% (w/v) were established for HY. Meanwhile, for CI, 13 levels between 0.03 and 0.15% (w/v) were determined for later validation steps. For each method (official and modified), the samples were analyzed randomly under intermediate precision conditions in three analytical batches (including different times and analysts) (Figure 2a). The random order was established using MS Excel (Microsoft, Redmond, WA, USA).

The false positive (FPR), FNR, sensitivity (SNR), selectivity (SLR) and reliability (RLR) rates were calculated by formulas based on contingency tables. RLR was estimated subtracting from 100 the FPR and FNR. Unreliability regions (URs) and detection limits (LOD) were calculated by logistic regression. The limits of the UR were the concentrations that corresponded to 0.05 and 0.95 probabilities of positive results. The LOD corresponded to the lower limit of the UR. The regression assumptions were verified, as proposed by Souza & Junqueira, applying the concepts proposed by Gondim et al., applying the concepts proposed by Gondim et al.

Robustness

The robustness was evaluated in a general full factorial design with two factors (reagents) and two levels (brands), resulting in four treatments. Two reagent brands of rosolic acid and ethyl alcohol were used to assess their influence on the detection of each acid neutralizer. Samples were spiked with the analytes at the lowest concentration levels that provided a 100% RLR in the previous validation step for both the official and modified methods. For samples adulterated with CA and HY, these concentrations corresponded to 0.025 and 0.0175% (w/v), respectively, for both methods. The concentrations evaluated for samples spiked with BI were 0.0425 and 0.0275% (w/v), considering the official and modified methods, respectively. For samples with CI added, the evaluated concentrations were 0.0175 and 0.009% (w/v), respectively, for the official and modified methods. Ten replicates were prepared per treatment. For each method, samples were randomly prepared and analyzed under repeatability conditions (Figure 2a). The methods were considered robust when the estimated RLRs were greater than or equal to 90%.

Complementary evaluation of the fitness for purpose

A study was carried out to verify the rosolic acid turnover range and to evaluate the relationship between the acidity and pH in milk samples. Raw milk samples with acidity levels at and above the allowed limits were spiked with acid neutralizers at two concentration levels, with three replicates per level. BI, CA and HY were added at concentrations 0.01 and 0.05% (w/v). CI was added at the levels 0.10 and 0.15% (w/v). These concentrations were selected based on the levels employed by Silva et al., but also in the concentration ranges evaluated in the previous single validation studies of the method. The acidity and pH of each sample were determined before and after addition of the acid neutralizer. Samples were also submitted to analysis by the rosolic acid qualitative methods (official and modified), and the visualized colors were compared with the respective RGB colors (Figure 2a).

Collaborative trial

The modified rosolic acid method was also validated by an interlaboratory process, as proposed by Gondim et al., with adaptations based on practical aspects. The collaborative trial included the following steps: i) sampling and test material preparation, ii) evaluation of the homogeneity and stability, iii) distribution of materials, iv) data analysis, and v) reporting of results (Figure 2b).

Sampling and test material preparation and evaluation

A raw milk sample (11 L) was obtained from the refrigerated storage tank of the Experimental Farm Professor Hélio Barbosa, EV/UFMG, and transported to the laboratory under refrigeration (2 to 7 °C) in a polypropylene gallon. The preparation of the test materials was performed when the raw milk sample reached an acidity level of approximately 0.19 mg L⁻¹, which was above that allowed by Brazilian legislation and occurred within approximately 24 hours after sample collection.

Then, the sample was divided into nine lots for the preparation of nine types of test materials. The first lot was used to prepare the test materials without the addition of an acid neutralizer (negative materials). The other lots were spiked with aqueous solutions of each acid neutralizer to obtain two concentration levels of each compound, totaling eight lots (types) of adulterated test materials (positive materials) (Figure 1b). These adulterated materials were added with 0.0175 and 0.0625% (w/v) BI, 0.0275 and 0.0625% (w/v) CA and 0.10 and 0.15% (w/v) CI. The concentration levels were the higher and lower concentrations that provided 100% RLRs in the previous single laboratory validation process.

Aliquots of 10 mL of each test material were packaged into 15 mL polypropylene conical tubes. The tubes were sealed, labeled with a random number code and kept under refrigeration (4 to 7 °C) until delivery to collaborators. To define the number of tubes to be prepared, it was considered, in addition to the quantity required to carry out the experiments (homogeneity, stability and collaborative trial), the possibility of the replacement of the test materials. The homogeneity tests were performed to study the variation between the test materials by the analysis of 10 test materials in duplicate. Stability tests involved two experimental designs: isochronous (10 test materials maintained at critical transportation conditions, 4 and 8 hours at (35 ± 5) °C, and analyzed in duplicate) and classical (10 test materials analyzed in duplicate at preparation and at the deadline for the analysis by collaborators). In both the homogeneity and stability tests, ten replicates of each type of test material were evaluated (Figure 2b).

Distribution of test materials

Eight laboratories representing different sectors of the dairy chain participated in the collaborative trial, with four laboratories from the public sector (three from governmental institutions and one linked to an university) and four being private, of which two were of first-party testing (from dairy industries) and the other two were of second-party testing (food analysis laboratories).

Each laboratory received 54 blind and coded test materials, totaling 486 results. Six replicates of each type of test material were sent: six of the negative test material and six replicates of each material containing each of the four acid neutralizers evaluated at two concentration levels, totaling 12 replicates per compound. The collaborators also received a non-blind blank sample corresponding to an unadulterated sample as a reference for results and the related documents (a material receipt form, the analytical method procedure and a form for the registration of analytical data) (Figure 2b).

The collaborators were instructed to analyze the test materials on the day and time previously agreed between all the laboratories. In addition, the materials should be analyzed by a single analyst, in a single analytical batch, following strictly the analytical procedure sent, and using the color scale as a complementary tool to classify the samples.

Data analysis and reports

The results received from the collaborators were submitted to an
evaluation to remove transcription or typing errors. Hierarchical cluster analysis (HCA) was performed to evaluate differences between the classes (productive sector) of laboratories that could influence the results of the interlaboratory validation. The results (positive or negative) obtained for each sample analyzed by the laboratory were used in this analysis. Comments and observations reported by the laboratories were also evaluated.

In the interlaboratorial validation, the performance parameters including the FPR, FNR, RLR, ACC and CON were evaluated, in addition to the estimation of the prediction intervals for the probability of detection. For each concentration of adulterant, the rates, ACC values and CON values were calculated, and they were considered satisfactory values if they corresponded to one false result. A false result can be a false positive or a false negative result, depends if the sample is negative (blank) or positive (spiked with acid neutralizer), respectively (Figure 2b). Six replicates of each material were distributed to each laboratory. Therefore, to be considered satisfactory, the estimated values for FPR and FNR should be less than or equal to 16.7%. For RLR, ACC and CON, the values should be bigger than or equal to 83.3%, 0.67 and 0.63, respectively.

RESULTS AND DISCUSSION

Optimization

The official method of rosolic acid establishes the addition of neutralized ethyl alcohol without specifying the concentration of the solution to be used. In these preliminary studies, was used an absolute ethyl alcohol (98%), which resulted in great difficulty in visualizing the results and high false result rates (FPR and FNR ≥ 50%).

The milk thermal stability can be tested by the alcohol test, which is based on the principle of the destabilization of the casein micelles by an ethyl alcohol solution and the observation of coagulation due to high acidity or salt imbalance. The test basically consists of the addition of ethyl alcohol solution to the milk. The concentration of the ethyl alcohol solution depends on the heat treatment to be applied to the milk and the shelf-life expected for the product to be obtained. NI 68/2006 establishes concentrations between 68 and 80%, while NI 62/2011 set the concentration of 72% to be used for the daily quality control of the raw refrigerated milk in the industrial establishment. Regardless of the concentration established by the NI, milk samples with high acidity should present a positive result for this test.

Milk samples with acidity at a normal level (between 0.14 and 0.18 mg L⁻¹), named non-acidic unstable milk, could present positive results for the alcohol test. Studies have shown that milk coagulation can be affected by the concentration of salt available in the medium, and the addition of small amounts of calcium or magnesium in milk makes the alcohol test positive, while citrate and phosphate cause the opposite effect. Other factors that can also affect milk stability are the season, diet and lactation stage. To minimize the coagulation effects and to facilitate the visualization of the results, a neutralized 68% ethyl alcohol solution in the later stages was used. Also, a second manual shaking step was included in the analytical procedure of the modified method, allowing the visualization of a more homogeneous staining by dissolving the rosolic acid solution in the sample.

Finally, to improve its sensitivity, a new criterion for positive results was suggested, which was based on the color scale previously presented: observation of a color different from that obtained for a non-blind blank sample (sample without adulteration with acidity at a standard level). In Figure 3a, the visualized colors are displayed as well as their classification in the RGB system.

Thus, the analytical procedure of the modified method consisted of adding 10 mL of ethyl alcohol (68%) neutralized to a 5 mL of milk sample, followed by manual shaking and addition of two drops of rosolic acid 2% in ethyl alcohol. After a second stage of manual shaking, the color observed was compared with that obtained for a non-blind blank sample previously submitted to the same procedure (Figure 1).

Single laboratory validation

Evaluation of rates, unreliability region, detection limit, accordance and concordance

The estimated values of rates, ACC and CON for each adulterant and concentration can be found in Table 1.

For both methods, the analysis of blank samples resulted in a SLR and RLR of 100% and a zero FPR, indicating the selectivity of the rosolic acid method. In samples with BI added, considering the official method, a 0% FNR was obtained from 0.0425% BI. For the modified method, a first 0% FNR was estimated at 0.0275% BI, although, at higher concentrations, FNRs between 10 and 3.3% were still observed. In the analysis of samples containing CA, the official method provided a 0% FNR from 0.0225% CA. However, for the modified method, a 0% FNR was estimated from 0.0250% CA, with the exception being at 0.030% CA, with a FNR of 3.3%. FNR values of 0% were obtained at levels greater than or equal to 0.0175% for samples with HY added in both methods. Regarding CI, in the official method, 0% FNRs were observed from 0.09 and 0.08% CI, respectively, for the official and modified methods. However, at higher levels, a 3.3% FNR was also obtained. The occurrence of false negatives results at high analyte concentrations could be a consequence of failures in sample preparation, analytical procedure or interpretation of the results by the analyst.

Silva et al. evaluated milk samples adulterated with HY at 0.01, 0.025, 0.05, and 0.1% by the rosolic acid method. In this study, only samples with 0.1% HY were classified as positive. A better performance was observed in the validation of the official method, in which, for 0.01% and levels above 0.02%, 80 and 100% positive results were observed, respectively. In the modified method, for the concentration of 0.01%, the rate of positive results was even higher at 93.3%, indicating the performance improvement of this version.

Mid-infrared spectroscopy (MIR) methods with different multivariate classification techniques have been used to identify raw milk adulteration by the addition of acid neutralizers, such as the work reported by Cassoli et al. In the experiment performed by these authors, all samples with 0.05% BI added were correctly identified as adulterated and 93.9% of samples adulterated with 0.075% CI were correctly classified. MIR was also evaluated with the soft independent modelling of class analogy (SIMCA) technique in a sequential strategy to detect adulterants in milk, although at higher levels. The MIR/SIMCA method was considered appropriate to detect samples added with 0.4% CA and 0.65% CI but not suitable to detect milk adulterated with BI and HY at 0.4%. Comparing these works with our study, the results were similar or better results. However, here the analytical technique employed was simpler and did not require the use of expensive equipment or software.

Considering samples with BI, ACC and CON values lower than 0.8 were estimated for both versions of the method, and these results can be attributed mainly to the performance of the first analytical batch (Table 1). Once more, these results could be attributed to failures in sample preparation, analytical procedure or interpretation of the results by the analyst, which were becoming suitable in the following analytical batches.

For the modified method, the performance of the first analytical
Table 1. False positive rate (FPR), false negative rate (FNR), selectivity rate (SLR), sensitivity rate (SNR), reliability rate (RLR), accordance (ACC) and concordance (CON) values obtained for the acid neutralizers in the single laboratory validation process.

| Sodium bicarbonate (% w/v) | Official method  | Modified method  |
|----------------------------|------------------|------------------|
|                            | FPR* or FNR | RLR or SLR* or RLR or SNR | ACC (analytical batch) | CON | FPR* or FNR | RLR or SLR* or RLR or SNR | ACC (analytical batch) | CON |
| 0.0000                     | 100.0*         | 100.0*             | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.0025                     | 100.0          | 0.0               | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.0050                     | 90.0           | 10.0              | 0.8 | 0.8 | 0.6 | 0.8 | 0.8 | 1.0 |
| 0.0075                     | 83.3           | 16.7              | 0.5 | 0.8 | 1.0 | 0.5 | 1.0 | 1.0 |
| 0.0100                     | 83.3           | 16.7              | 0.6 | 0.5 | 1.0 | 0.6 | 1.0 | 1.0 |
| 0.0125                     | 73.3           | 26.7              | 0.5 | 0.4 | 1.0 | 0.5 | 1.0 | 0.8 |
| 0.0150                     | 60.0           | 40.0              | 0.4 | 0.4 | 0.6 | 0.4 | 0.8 | 0.9 |
| 0.0175                     | 33.3           | 66.7              | 0.5 | 0.6 | 0.5 | 0.5 | 0.6 | 0.7 |
| 0.0200                     | 16.7           | 83.3              | 1.0 | 0.8 | 0.5 | 1.0 | 0.8 | 0.9 |
| 0.0225                     | 23.3           | 76.7              | 0.8 | 0.6 | 0.5 | 0.8 | 0.8 | 0.9 |
| 0.0250                     | 20.0           | 80.0              | 1.0 | 0.8 | 0.4 | 1.0 | 1.0 | 1.0 |
| 0.0275                     | 10.0           | 90.0              | 1.0 | 1.0 | 0.5 | 1.0 | 1.0 | 1.0 |
| 0.0300                     | 3.3            | 96.7              | 1.0 | 0.8 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.0325                     | 10.0           | 90.0              | 0.5 | 1.0 | 1.0 | 0.5 | 0.6 | 0.7 |
| 0.0350                     | 10.0           | 90.0              | 0.6 | 0.8 | 1.0 | 0.6 | 0.8 | 0.9 |
| 0.0375                     | 10.0           | 90.0              | 0.5 | 1.0 | 0.5 | 0.5 | 0.6 | 0.7 |
| 0.0400                     | 3.3            | 96.7              | 1.0 | 0.8 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.0425                     | 0.0            | 100.0             | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.0450                     | 0.0            | 100.0             | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.0475                     | 0.0            | 100.0             | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 0.0500                     | 0.0            | 100.0             | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
Table 1. False positive rate (FPR), false negative rate (FNR), selectivity rate (SLR), sensitivity rate (SNR), reliability rate (RLR), accuracy (ACC) and concordance (CON) values obtained for the acid neutralizers in the single laboratory validation process (cont.)

| Sodium carbonate (%) | Official method | Modified method |
|----------------------|----------------|-----------------|
|                      | FPR* or FNR   | ACC (analytical batch) | CON | FPR* or FNR | ACC (analytical batch) | CON |
| 0.0000               | 100.0*        | 1.0 1.0 1.0     | 1.0 | 100.0*      | 1.0 1.0 1.0     | 1.0 |
| 0.0025               | 100.0         | 1.0 1.0 1.0     | 1.0 | 100.0       | 1.0 1.0 1.0     | 1.0 |
| 0.0050               | 96.7          | 1.0 1.0 0.8     | 0.9 | 100.0       | 1.0 1.0 1.0     | 1.0 |
| 0.0075               | 83.3          | 1.0 1.0 0.4     | 0.7 | 76.7        | 23.3 0.5 0.6    | 0.6 |
| 0.0100               | 80.0          | 0.8 0.6 0.5     | 0.7 | 56.7        | 43.3 0.5 0.5    | 0.5 |
| 0.0125               | 73.3          | 0.5 0.6 0.5     | 0.6 | 40.0        | 60.0 0.4 0.5    | 0.5 |
| 0.0150               | 40.0          | 0.8 0.8 0.6     | 0.4 | 26.7        | 73.3 1.0 0.5    | 0.6 |
| 0.0175               | 16.7          | 0.8 0.8 0.5     | 0.7 | 6.7         | 93.3 1.0 0.6    | 0.9 |
| 0.0200               | 10.0          | 0.8 1.0 0.6     | 0.8 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.0225               | 6.7           | 0.8 0.8 1.0     | 0.9 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.0250               | 6.7           | 0.8 0.8 1.0     | 0.9 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.0275               | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |
| 0.0300               | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |
| 0.0325               | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |
| 0.0350               | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |
| 0.0375               | 3.3           | 1.0 1.0 0.8     | 0.9 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.0400 to 0.0500     | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |

Table 1. False positive rate (FPR), false negative rate (FNR), selectivity rate (SLR), sensitivity rate (SNR), reliability rate (RLR), accuracy (ACC) and concordance (CON) values obtained for the acid neutralizers in the single laboratory validation process (cont.)

| Sodium hydroxide (%) | Official method | Modified method |
|----------------------|----------------|-----------------|
|                      | FPR* or FNR   | ACC (analytical batch) | CON | FPR* or FNR | ACC (analytical batch) | CON |
| 0.0000               | 100.0*        | 1.0 1.0 1.0     | 1.0 | 100.0*      | 1.0 1.0 1.0     | 1.0 |
| 0.0025               | 100.0         | 1.0 1.0 1.0     | 1.0 | 100.0       | 1.0 1.0 1.0     | 1.0 |
| 0.0050               | 80.0          | 0.4 1.0 0.8     | 0.6 | 70.0        | 30.0 0.5 0.6    | 0.6 |
| 0.0075               | 63.3          | 0.4 0.8 0.4     | 0.5 | 20.0        | 80.0 0.6 0.5    | 0.7 |
| 0.0100               | 20.0          | 0.8 0.5 0.8     | 0.7 | 6.7         | 93.3 1.0 0.6    | 0.9 |
| 0.0125               | 10.0          | 1.0 0.6 0.8     | 0.8 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.0150               | 10.0          | 1.0 0.6 0.8     | 0.8 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.0175 to 0.0275     | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |

Table 1. False positive rate (FPR), false negative rate (FNR), selectivity rate (SLR), sensitivity rate (SNR), reliability rate (RLR), accuracy (ACC) and concordance (CON) values obtained for the acid neutralizers in the single laboratory validation process (cont.)

| Sodium citrate (%) | Official method | Modified method |
|--------------------|----------------|-----------------|
|                    | FPR* or FNR   | ACC (analytical batch) | CON | FPR* or FNR | ACC (analytical batch) | CON |
| 0                   | 100.0*        | 1.0 1.0 1.0     | 1.0 | 100.0*      | 1.0 1.0 1.0     | 1.0 |
| 0.03                | 80.0          | 1.0 0.5 0.6     | 0.7 | 70.0        | 30.0 0.4 0.5    | 0.6 |
| 0.04                | 60.0          | 0.8 0.5 0.4     | 0.5 | 40.0        | 60.0 0.8 0.6    | 0.4 |
| 0.05                | 66.7          | 0.5 0.5 0.8     | 0.5 | 33.3        | 66.7 0.5 0.8    | 0.5 |
| 0.06                | 26.7          | 0.5 0.8 0.6     | 0.6 | 6.7         | 93.3 0.8 0.8    | 0.9 |
| 0.07                | 13.3          | 0.5 1.0 0.8     | 0.8 | 6.7         | 93.3 0.8 0.8    | 0.9 |
| 0.08                | 13.3          | 0.5 1.0 1.0     | 0.7 | 0.0         | 100.0 1.0 1.0   | 1.0 |
| 0.09                | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |
| 0.10                | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |
| 0.11                | 3.3           | 1.0 1.0 0.8     | 0.9 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.12                | 3.3           | 1.0 1.0 0.8     | 0.9 | 3.3         | 96.7 1.0 0.8    | 0.9 |
| 0.13 and 0.15       | 0.0           | 1.0 1.0 1.0     | 1.0 | 0.0         | 100.0 1.0 1.0   | 1.0 |

*FPR and SLR were estimated only for unadulterated samples, i.e., 0% acid neutralizer. ACC and CON values in bold correspond to levels outside the unreliability region (UR).

Conducted according to premises and guidelines described in the main international guides and that, despite the lack of a specific guide for the validation of multivariate methods, methodologies described for univariate methods have been adapted for an adequate evaluation of the performance of multivariate methods.

Robustness
In the robustness evaluation, it was considered that the reagent brands did not influence the method performance, since a 100% RLR was obtained in all experiments, considering the four acid neutralizers evaluated and both the official and modified methods.
Complementary evaluation of the fitness for purpose

In Figure 3b, the color scale for the rosolic acid method is presented as well as the acidity values, expressed in mg L\(^{-1}\) of lactic acid, and the pH associated with the visualized colors.

Before the addition of neutralizers, samples with a normal acidity index (approximately 0.15 mg L\(^{-1}\)) and those with an acidity level above the allowed limit (mean of 0.20 mg L\(^{-1}\)) presented negative results for the rosolic acid test, i.e., a salmon color. However, those samples with an acidity above that allowed by the legislation showed a less intense salmon color. After the addition of the neutralizers, milk samples with an acidity above the established limit of 0.18 mg L\(^{-1}\) (with a mean of 0.20 mg L\(^{-1}\)) presented acidity values within the legislation limits. The pH values of these samples varied between 6.2 and 6.9 and were the highest values obtained after the addition of the neutralizers. The rosolic acid method, official and modified versions, provided positive results (light pink) for samples with HY added and negative for those with other neutralizers added.

As expected, samples with satisfactory initial acidity (0.14-0.18 mg L\(^{-1}\)) presented positive results after the addition of the neutralizing compounds (from light pink to deep pink). The new acidity values ranged from 0.14 to 0.07 mg L\(^{-1}\), with higher values obtained for samples adulterated with BI and CI and smaller acidity values for those with HY added. In relation to the pH, a greater variation was observed for samples that contained HY, with values between 6.8 and 7.4. In addition, those added to the other adulterants had a slight change in pH between 6.8 and 7.0. The observed behavior for the acid neutralizers can be explain by the nature of these compounds. Strong bases, such as HY, can alter the pH and the acidity of the medium more significantly than weak bases (such as CI and BI).

The results indicated the ability of the method to detect neutralizers in samples. However, for samples that had high acidity values (above 0.18 mg L\(^{-1}\)), a method limitation was observed in detecting the addition of adulterants classified as weak bases. The method was suitable for the adulterant HY but limited for the other adulterants tested that provide positive and false negatives results. Again, this can be explained by the nature of the bases and could happen because the final pH was under 7.0 and, therefore, the indicator did not turn its color. In a similar study with the phenolphthalein method for the detection of acid neutralizers, Silva et al.\(^{10}\) observed that the method was also not able to detect the presence of HY in milk samples when a precise neutralization was performed. According to these authors, this was due to the sensitivity and selectivity of the indicator used. In fact, phenolphthalein has a greater turning range than rosolic acid, between 8.2 and 10.0.\(^{3}\)

Collaborative trial

Homogeneity and stability tests

In the homogeneity evaluation, the analysis of the nine lots of test materials resulted in a 100% RLR, i.e., false results were not obtained. Negative results were obtained for all of the 10 negative materials that were analysed in duplicate and positive results were obtained for all of the 10 positive materials of each acid neutralizer in each concentration level that were analysed in duplicate.

The short-term stabilities of the test materials have been demonstrated. For the isochronous design simulating transport conditions, 90% SLRs were obtained for the test materials with 0.0425% BI, 0.10% CI and 0.0275% CA added and maintained at temperatures between 30 and 40 °C for 10 hours. However, according to the established criteria (RLR ≥ 90%), the results could not be associated with the temperature conditions to which the materials were subjected.

Hierarchical cluster analysis (HCA)

Two clusters could be observed in the dendrogram shown in Figure 5. One of these was composed of laboratories 6 and 2, representatives of official laboratories and dairy industries, respectively. The second group was formed by laboratories of all the segments involved in the collaborative trials. Thus, the HCA indicated that the classes to which the laboratory collaborators belonged did not influence the results obtained in the interlaboratory validation.

Collaborative trial

Satisfactory results were observed for materials not adulterated,
Detection of acid neutralizers in fraudulent milk: full validation of a classical qualitative method

with FPRs ≤ 16.7% (or RLRs ≥ 83.3%) for all collaborators. The exception was the Laboratory 7, which observed a 33.3% FPR (or 66.7% RLR) (Table 2). Considering the sum of the results of all laboratories, the RLR was 97.2%, indicating the selectivity of the method.

Still analysing the overall results for the adulterated materials, FNRs between 2.1 and 29.8% were observed. Satisfactory results of the RLR (≥ 83.3%) were obtained only for the materials adulterated with the highest concentration of each neutralizer. Individually, Laboratories 2 and 6 presented worse performance in relation to the FNR (Table 2). Similar behavior was observed for the ACC values, with Laboratory 2 presenting a poor performance. Unsatisfactory CON values (less than 0.63) were obtained just for test materials adulterated with 0.0425% BI and 0.0275% CA. (Table 3). These unsuitable results could be attributed to the lack or inadequate use of the non-blind blank sample and color standards. In fact, Laboratory 2 reported difficulties in comparing the colors with the non-blind blank sample and consequently identifying a positive result, as presented in the section 3.3.4 (Comment of collaborators). However, problems were not reported by Laboratory 6.

In addition, it is important to highlight that the difference in the acidity levels of the samples used in this process was different those employed in the single laboratory validation. In the collaborative trial, samples had acidity above the legislation limit, which did not always occur in single laboratory validation. Thus, the difference in the acidity of samples could explain the inferior performance obtained in the collaborative trial, since the complementary evaluation of the fitness for purpose indicated a method limitation in detecting adulterated samples with high levels of acidity.

Table 4 shows the prediction intervals for the probability of detection as well as the statistical parameters used in the calculation of these intervals. Statistical calculations, performed according to Macarthur and Holst’s protocol, corroborate the results presented by the estimates of rates, ACC and CON values. The probability of obtaining false positive results was calculated as 0.06, confirming the selectivity of the method. For materials adulterated with HY and CI, the probability of obtaining false negative results was calculated as 0.08.

Table 2. False negative (FNR), false positive (FPR) and reliability (RLR) rates estimated in the collaborative trial (interlaboratorial validation process)

| Laboratory | 0.0 | Sodium hydroxide (%., w/v) | Sodium bicarbonate (%., w/v) | Sodium carbonate (%., w/v) | Sodium citrate (%., w/v) |
|------------|-----|-----------------------------|-------------------------------|---------------------------|-------------------------|
|            |     | 0.0175                      | 0.0625                        | 0.0425                    | 0.0625                  |
| 1          | 0.0 | 100.00                      | 100.00                        | 16.7                      | 83.3                    |
| 2          | 0.0 | 100.00                      | 50.0                          | 50.0                      | 0.0                     |
| 3          | 0.0 | 100.00                      | 0.0                           | 0.0                       | 0.0                     |
| 4          | 0.0 | 100.00                      | 0.0                           | 0.0                       | 0.0                     |
| 5          | 16.7| 83.3                        | 0.0                           | 0.0                       | 0.0                     |
| 6          | 0.0 | 100.00                      | 0.0                           | 0.0                       | 0.0                     |
| 7          | 33.3| 66.7                        | 0.0                           | 0.0                       | 0.0                     |
| 8          | 0.0 | 100.00                      | 0.0                           | 0.0                       | 0.0                     |
| Total      | 6.3 | 93.8                        | 20.8                          | 21.1                      | 97.9                    |

RLR values in bold indicate results lower than 83.3%, i.e., unsatisfactory.

Table 3. Accordance (ACC) and concordance (CON) values estimated in the collaborative trial (interlaboratorial validation process)

| Laboratory | 0.0 | Sodium hydroxide (%., w/v) | Sodium bicarbonate (%., w/v) | Sodium carbonate (%., w/v) | Sodium citrate (%., w/v) |
|------------|-----|-----------------------------|-------------------------------|---------------------------|-------------------------|
|            |     | 0.0175                      | 0.0625                        | 0.0425                    | 0.0625                  |
| 1          | 1.00| 0.67                        | 1.00                          | 0.67                      | 1.00                    |
| 2          | 1.00| 0.40                        | 1.00                          | 1.00                      | 0.40                    |
| 3          | 1.00| 1.00                        | 1.00                          | 1.00                      | 1.00                    |
| 4          | 1.00| 1.00                        | 1.00                          | 1.00                      | 1.00                    |
| 5          | 0.67| 1.00                        | 1.00                          | 1.00                      | 1.00                    |
| 6          | 1.00| 0.47                        | 1.00                          | 0.47                      | 1.00                    |
| 7          | 0.47| 1.00                        | 1.00                          | 0.67                      | 1.00                    |
| 8          | 1.00| 1.00                        | 1.00                          | 1.00                      | 1.00                    |

CON values in bold indicate unsatisfactory results for ACC and CON, i.e., values lower than 0.67 and 0.63, respectively.
Table 4. Statistical parameters as well as lower (5%) and upper (95%) limits estimated for the probability of detection of acid neutralizers in milk

| Acid neutralizer  | Concentration (% w/v) | N  | X  | \( p \)     | \( s_p \)     | Lower limit | Upper limit |
|-------------------|------------------------|----|----|------------|--------------|-------------|-------------|
| -                 |                        | 0  | 48 | 37         | 0.063        | 0.0155      | 0.0392      | 0.1398      |
| Sodium hydroxide  | 0.0175                 | 48 | 38 | 37         | 0.792        | 0.0456      | 0.6839      | 0.8742      |
| Sodium bicarbonate| 0.0625                 | 48 | 47 | 34         | 0.979        | 0.0074      | 0.9214      | 0.9963      |
| Sodium carbonate  | 0.0425                 | 48 | 41 | 34         | 0.708        | 0.0554      | 0.5932      | 0.8054      |
| Sodium citrate    | 0.0275                 | 48 | 41 | 34         | 0.854        | 0.0304      | 0.7557      | 0.9220      |
| Sodium citrate    | 0.0625                 | 48 | 43 | 34         | 0.896        | 0.0368      | 0.8063      | 0.9512      |

\( N \) = total number of analyses; \( X \) = total number of positive results; \( p \) = estimated mean probability of detection; \( s_p \) = standard deviation of estimates of the probability of detection from the individual laboratories (Macarthur and Holst \(^{14}\)).

the detection probabilities were higher than 0.80, while for the tests with CA and BI added, the probabilities of detection were estimated between 0.708 and 0.896. These probability values indicated a greater sensitivity of the method to the higher levels of acid neutralizers. Regarding the variability of the results presented between the laboratories, that is, the intermediate precision, a significant variation in the prediction intervals is observed for the probability of detection of the neutralizers.

**Acknowledgements**

The authors acknowledge the experimental farm Professor Hélio Barbosa of the Escola de Veterinária (EV/UFMG) for providing the raw milk samples. The authors also express their appreciation to the laboratory collaborators: GMO-Centro de Pesquisas e Controle de Qualidade Ltda.; Ita Alimentos-Laticínios Ita Indústria e Comércio de Alimentos Ltda.; Iambé Alimentos S.A.; Laboratório de Análise Físico-Química em Alimentos do Laboratório de Química Agropecuária do Instituto Mineiro de Agropecuária (LAFQ/LQA/IMA); Laboratório de Química Bromatológica da Fundação Ezequiel Dias (FUNED); Laboratório Nacional Agropecuário de Minas Gerais do Ministério da Agricultura, Pecuária e Abastecimento (LANAGRO-MG/MAPA); Laboratório de Bromatología – Unidade de Pesquisa Análise de Alimentos da Faculdade de Farmácia da Universidade Federal de Minas Gerais (BRO-UPAA/FAFAR/UFMG) - and Núcleo Global de Análise e Pesquisa Ltda. (NUGAP).

**References**

1. Handford, C. E.; Campbell, K.; Elliott, C. T.; *Compr Rev Food Sci. Food Saf.* 2016, 15, 130.
2. www.agencia.cnptia.embrapa.br/Agencia8/AG01/arvore/AG01_195_21720039246.html, acessada em janeiro 2021.
3. Troncoso, V. M.; *Manual para inspeção da qualidade do leite*, 4th ed., Editora da UFSC, Santa Maria, 2010.
4. Sherbon, J. S.; *Em Fundamentals of Dairy Chemistr*- Physical properties of Milk, Wong, N. F., ed.; Aspen Publishers: New York, 1998.
5. Santos, P. M.; Pereira-Filho, E. R.; Rodriguez-Saona, L. E.; *Food Chem.* 2013, 138, 19.
6. Sharma, K.; Paradakar, M.; *Food Security* 2010, 2, 97.
7. Brasil - Ministério da Agricultura, Pecuária e Abastecimento, Instrução Normativa nº 62 de 29 de dezembro, 2011.
8. Ministry of Health and Family Welfare of India; *Manual of Methods of Analysis of Foods - Milk & Milk Products*, 2015.
9. Brasil - Ministério da Agricultura, Pecuária e Abastecimento, Instrução Normativa nº 68 de 12 de dezembro, 2006.
10. Silva, L. C. D.; Tamanini, R.; Pereira, J. R.; Rios, E. A.; Ribeiro-Junior, J. R.; Beloti, V.; *Cienc. Rural* 2015, 45, 1613.
11. Ritter, H. L.; *Introducción a la Química*, Editorial Reverté S. A.: Barcelona, 1956.
12. State Council of the People’s Republic of China, Regulation on the Supervision and Administration of the Quality and Safety of Dairy Products – Order of the State Council of the People’s Republic of China (No. 536), 2008.
13. Trullols, E.; Ruisánchez, I.; Rius, F. X.; *Trend Anal. Chem.* 2004, 23, 137.
14. Macarthur, R.; Holst, C.; *Anal. Methods* 2012, 4, 2744.
15. AOAC International; *J. AOAC Int.* 2014, 97, 1492.
16. Gondim, C. S.; Sousa, R. C. S.; Palhares, M. P. P.; Junqueira, R. G.; Souza, S. V. C.; *Anal. Methods* 2015, 7, 9692.
17. Gondim, C. S.; Junqueira, R. G.; Souza, S. V. C.; *Food Analytical Methods* 2016, 9, 2509.
18. Gondim, C. S.; Coelho, O. A. M.; Alvarenga, R. L.; Junqueira, R. G.; Souza, S. V. C.; *Anal. Chim. Acta* 2014, 830, 11.
19. Krause, J.; *Color Index*, How Book: Cincinnati, 2010.
20. Souza, S. V. C.; Junqueira, R. G.; *Anal. Chim. Acta* 2005, 552, 25.
21. Langton, S. D.; Chevennement, R.; Nagelkeke, N.; Lombard, B.; *Int. J. Food Microbiol.* 2002, 79, 171.
22. Thompson, M.; Ellison, S. L. R.; Wood, R.; *Pure Appl. Chem.* 2006, 78, 145.
23. Ellison, S. L. R.; Fearn, T.; Trends Anal. Chem. 2005, 24, 468.
24. Ward, J. H.; J. Am. Stat. Assoc. 1963, 58, 236.
25. Zanela, M. B.; Fischer, V. Em Qualidade do leite bovino: variações no trópico e no subtrópico; González, F. D., Pinto, A. T., Zanela, M. B., Fischer, V., Bondan, C., eds.; UPF Editora: Passo Fundo, 2011.
26. Cassoli, L. D.; Sartori, B.; Machado, P. F.; Rev. Bras. Zootec. 2011, 40, 2591.
27. Gondim, C. S.; Junqueira, R. G.; Souza, S. V. C., Ruisánchez, I.; Callao, M. P.; Food Chem. 2017, 230, 68.
28. EC - European Commission. Comission Regulation No. 2002/657/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results; Off. J. Eur. Communities 2002, 8.
29. EURACHEM; The fitness for purpose of analytical methods, a laboratory guide to method validation and related topics, 2nd ed., LGC: Teddington, 2014.
30. Thompson, M.; Ellison, S. L. R.; Wood, R.; Pure Appl. Chem. 2002, 74, 835.
31. Botelho, B. G.; Mendes, B. A. P.; Sena, M. M.; Quim. Nova 2013, 36, 1416.
32. Botelho, B. G.; Reis, N.; Oliveira, L. R.; Sena, M. M.; Food Chem. 2015, 181, 31.