cosmetics

MDPI

Article

Evaluating Laboratory Performance in a Proficiency Test for Antimicrobial Protection of a Cosmetic Formulation

Sabrina Bouakache [1,*], Thanh Thuy Nguyen [1], Margaux François Elissalde [2], Anne Tirard [1] and Abdelkader Boubetra [1]

[1] Bureau Interprofessionnel d'Études Analytiques (BIPEA), 189 Rue d'Aubervilliers, 75018 Paris, France; tnguyen@bipea.org (T.T.N.); atirard@bipea.org (A.T.); aboubetra@bipea.org (A.B.)
[2] L'OREAL, 100 Avenue de Stalingrad, 94550 Chevilly-Larue, France; margaux.elissalde@loreal.com
* Correspondence: sbouakache@bipea.org; Tel.: +33-1-75-43-17-63

Abstract

Accurate assessments of antimicrobial protection of cosmetic products play a crucial role in monitoring their quality and user health. The reliability of these analyses depends on the method used and the competence of cosmetic laboratories. This paper presents the design and results of a proficiency test organized in November 2023 to assess the performance of 37 laboratories participating in the determination of the logarithmic reduction in *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Candida albicans*, and *Aspergillus brasiliensis* 7 days after inoculation in a cosmetic formulation. Participating laboratories performed their duplicate analysis on each test microorganism under repeatability conditions, using the same analytical method. In order to evaluate the consistency of laboratory performance, two trials were scheduled at a one-week interval, and they were conducted by the same analyst, applying the same analytical procedure on identical samples. The results of this study demonstrated the overall good application of the method in the determination of microbial reduction by participating laboratories and highlighted the robustness of integrating method precision assessment into proficiency testing for monitoring the analytical performance of cosmetic microbiology laboratories.

Keywords: proficiency testing; cosmetic microbiology; antimicrobial efficacy testing; method precision; quality control

1. Introduction

Cosmetic products are complex matrices. The presence of nutrient substances such as lipids, proteins, polysaccharides, glucosides, amino acids, vitamins, sterols, and peptides can favor the viability and growth of microorganisms in cosmetic products [1]. The excessive concentration of microorganisms and the presence of certain pathogenic microorganisms can alter the physicochemical properties of a cosmetic product and pose safety risks to users [2–6]. Therefore, the microbiological quality of a product launched on the market should be monitored, and its microbiological risks should be assessed to ensure that the microorganisms developed during the use of a product have no undesirable impact on its quality and user health [7].

The risks of microbiological contamination in the course of product use depend on many factors, such as its composition and the microbial growth conditions, including temperature, pH, water activity, nutrient content, and preservatives. According to the ISO 29621 standard, low-risk products have water activity values ≤ 0.75, pH ≤ 3 or ≥10, and alcohol

Check for updates

Academic Editor: Enzo Berardesca

Received: 8 October 2026
Revised: 26 October 2026
Accepted: 27 October 2026
Published: 30 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

content ≥20% [7]. To inhibit the growth of microorganisms, natural or synthetic chemical preservatives are frequently added to cosmetic products by manufacturers [8,9]. However, reliable assessment of their antimicrobial preservation efficacy remains challenging due to the intrinsic heterogeneity in microbial distribution, the complexity of cosmetic matrices, and operational variability in method implementation.

The method used to assess the antimicrobial protection efficacy of a cosmetic product is standardized [10] so that a laboratory's result is comparable with the results obtained by other laboratories applying the same method. Although crucial, the application of standard methods alone does not guarantee the reliability of analytical results. Cosmetic laboratories can demonstrate the reliability of their results by participating in proficiency tests (PTs), which is a mandatory requirement for ISO 17025 accreditation [11]. PTs involve the analysis of homogeneous and stable samples by different laboratories, with results evaluated against predefined criteria. While proficiency testing programs are designed to assess the results of laboratories in a single trial, the assessment of individual laboratory performance consistency over time remains an understudied aspect of quality assurance in cosmetic analysis.

This paper describes the design and results of a proficiency test organized by Bureau Interprofessionnel d'Études Analytiques (BIPEA) in November 2023 to assess the analytical performance of 37 laboratories participating in the assessment of the antimicrobial protection efficacy of a cosmetic formulation. All laboratories applied the same analytical method to determine the logarithmic reduction values of *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Candida albicans*, and *Aspergillus brasiliensis* 7 days after they were inoculated in the test cosmetic formulation. All participating laboratories performed two successive counts, enabling evaluation of their performance under repeatability and reproducibility conditions. In addition, in order to evaluate the consistency of laboratory performance over time, the second trial was conducted one week after the first trial, involving the same analysts applying the same testing procedure on identical samples. The results of the study highlighted the robustness of integrating method precision evaluation into a routine proficiency test for controlling and monitoring laboratory quality in cosmetic microbiological analysis.

2. Materials and Methods

This proficiency test was developed and implemented in compliance with the requirements of the ISO/IEC 17043 standard [12], including four principal stages: experimental design, sample preparation, laboratory analysis, and statistical analysis of the results.

2.1. Proficiency Test Design

The design of this PT was summarized in Table 1. The test procedure was adapted from the ISO 11930 standard [10]. As the cosmetic formulation selected for the PT was of low antimicrobial protection, the experiment was conducted over a period of 7 days only. The PT was described in the main steps below:

- The cosmetic formulation was produced and packaged in 250 g glass containers and sent to the participating laboratories for analysis on the specified date.
- The laboratories inoculated the cosmetic formulation samples with a suspension of five strains, namely *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Candida albicans*, and *Aspergillus brasiliensis*, prepared in accordance with ISO 11930, for the whole duration of the test.
- In order to ensure that the inocula were well homogenized, the initial concentration of microorganisms after one minute of contact between the inoculated suspension and

the cosmetic formulation was determined for each strain. It varied from 1×10^6 to 3×10^6 CFU/g for all test microorganisms in this PT.

- The number of surviving microorganisms was analyzed in duplicate and determined at the defined interval of 7 days for all strains under repeatability conditions. For each test microorganism, the difference between the initial concentration (N_0) and the count at day 7 (N_7) was calculated and expressed in $\log_{10}$ CFU/g to determine the microbial reduction value. This value was compared to the evaluation criteria of the ISO 11930 standard at 7 days.
- In order to assess the consistency of test results, two trials were scheduled one week apart on identical samples and conducted by the same analyst, using the same strains, media, and equipment, following the same analytical procedure.

Table 1. Experimental design.

Test Microorganism	Initial Concentration (CFU/g)	Results of Microbial Reduction ($\log_{10}$ CFU/g)			
		Trial 1 (Day 7 of Week 1)		Trial 2 (Day 7 of Week 2)	
		Replicate 1	Replicate 2	Replicate 1	Replicate 2
Escherichia coli	1×10^6–3×10^6	✓ *	✓	✓	✓
Pseudomonas aeruginosa	1×10^6–3×10^6	✓	✓	✓	✓
Staphylococcus aureus	1×10^6–3×10^6	✓	✓	✓	✓
Candida albicans	1×10^6–3×10^6	✓	✓	✓	✓
Aspergillus brasiliensis	1×10^6–3×10^6	✓	✓	✓	✓

* ✓ indicates that microbial quantification was performed.

2.2. Sample Preparation

The product used for this PT was a hair mask formulated as an oil-in-water emulsion, produced and packaged in 250 g glass jars by an external manufacturer in compliance with the PT provider's quality management requirements. The microbiological quality of the product was controlled and guaranteed by the manufacturer. Each sample was labeled with a unique sample number and shipped to participating laboratories at ambient temperature on the specified date. Each laboratory received two samples for analysis.

2.3. Laboratory Analysis

Participating laboratories were required to analyze the samples received within the allotted time frame of four weeks. In order to aid the interpretation of the results, they were also asked to give information on the date of analysis, the strains used, the broths and agar media used for inoculation and enumeration, incubation duration, and incubation temperature.

Most laboratories used the ISO 11930 standard strains (ATCC®8739™ for *Escherichia coli*, ATCC®9027™3 or other equivalent national collection strains for *Pseudomonas aeruginosa*, ATCC®6538™ for *Staphylococcus aureus*, ATCC®10231™ for *Candida albicans*, and ATCC®16404™ or other equivalent national collection strains for *Aspergillus brasiliensis*). The inocula were prepared using tryptic soy agar (TSA) petri dishes. Eugon LT modified broth and TSA were used for microbial enumeration at 7 days. All laboratories incubated the test microorganisms at 32 °C ± 2. However, the incubation time varied from 48 to 96 h for bacteria, 48–144 h for *Candida albicans*, and 24–168 h for *Aspergillus brasiliensis*.

2.4. Statistical Analysis

The particular feature of this study was the integration of method precision evaluation into a proficiency test. The determination of microbial reduction values was based on the ISO 11930 standard [10]. The criteria for evaluation of individual laboratory performance were established according to the ISO 13528 standard [13]. Method precision, expressed as repeatability and reproducibility standard deviations, was estimated in accordance with the ISO 5725-2 and ISO 5725-5 standards [14,15].

2.4.1. Determination of Microbial Reduction Values

In microbiology, colony counts above 35 CFU per plate frequently conform to the lognormal distribution [16]. Results were transformed and expressed in $\log_{10}$ CFU/g to conform to an approximately normal distribution before statistical analysis [17].

In this study, a result submitted for proficiency assessment was the logarithmic reduction in *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Candida albicans*, and *Aspergillus brasiliensis* counts at day 7 after they were inoculated and calculated using Equation (1):

$$x_i = \log_{10}(N_{0,i}) - \log_{10}(N_{7,i}), \quad (1)$$

where i = laboratory i participating in the proficiency test, x_i = the logarithmic microbial reduction result of the laboratory i, N_0 = initial concentration (CFU/g), and N_7 = number of surviving microorganisms at day 7 (CFU/g). There was a reduction in the microbial population if $x_i > 0$. There might be no reduction if $x_i = 0$, and there might be an increase in the microbial population if $x_i < 0$. On day 7, the cosmetic formulation was considered adequately protected against microbial growth when the following criteria were met: $x_i \geq 3$ $\log_{10}$ CFU/g for *Escherichia coli*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus*; $x_i \geq 1$ $\log_{10}$ CFU/g for *Candida albicans;* and $x_i \geq 0$ $\log_{10}$ CFU/g for *Aspergillus brasiliensis* [10].

2.4.2. Determination of Assigned Values and Standard Deviations for Proficiency Assessment

Results were assessed in accordance with Annex C of the ISO 13528 standard [13]. Accordingly, a robust mean and its associated standard deviation (SD) were estimated by applying Algorithm A to the quantitative results reported by participating laboratories, and they were chosen as the consensus assigned value (x_{pt}) and the standard deviation for proficiency assessment (σ_{pt}), respectively. Algorithm A defined the initial robust mean as the median of the results and set the initial robust standard deviation at 1.483 times the median of the deviations from the initial mean. Through a process known as winsorization, the new robust mean and standard deviation were iteratively calculated by down-weighting extreme deviations rather than excluding outlying results until successive iterations produced no further change in the estimates. This robust statistical method is appropriate for approximately normally distributed data in which the proportion of outlying results does not exceed 20% [13]. In this PT, microorganisms were analyzed in duplicate, so the assigned value was estimated from the means of the replicate results of all participating laboratories.

2.4.3. Criteria for Evaluating Individual Laboratory Performance

A tolerance interval of 2 robust standard deviations from the assigned value ($x_{pt} \pm 2\sigma_{pt}$) was chosen as a performance assessment criterion (the coefficient of 2 corresponds approximately to 95% of the observed values under the assumption of the "normal" distribution of the results). The results falling in the tolerance interval were considered to be acceptable.

Performance can also be expressed in the form of a score so that results are comparable for different analytes and PT rounds. A standardized score frequently used to assess the

performance of individual laboratories in proficiency testing is a z-score [13]. In this PT, the z-score for a result x_i was calculated as

$$z_i = \frac{\left(x_i - x_{pt}\right)}{\sigma_{pt}}, \quad (2)$$

where i = laboratory i participating in the proficiency test, x_i = test result of laboratory i, x_{pt} = assigned value, σ_{pt} = standard deviation for proficiency assessment, and z_i = z-score calculated for the test result x_i. The result x_i was considered to be acceptable if $|z_i| \leq 2.0$, warning (or questionable) if $2.0 < |z_i| < 3.0$, and unacceptable if $|z_i| \geq 3.0$.

For each laboratory, the high number of acceptable results and the low number of unacceptable results are indicators of good performance. The results with $|z_i| > 2.0$ may be subject to an internal investigation to find out the sources of these deviations, and corrective or preventive actions can be implemented to improve the laboratory's performance.

2.4.4. Evaluation of Method Precision

Method precision is commonly characterized by repeatability and reproducibility and expressed as standard deviations of replicate results analyzed under repeatability and reproducibility conditions. In order to evaluate the precision of the duplicate results obtained under repeatability conditions (in the same laboratory, by the same analyst, using the same strains, media, and equipment, applying the same test procedure, within a short time period) and under reproducibility conditions (involving different laboratories, analysts, strains, media, equipment, and test procedures), robust estimates of repeatability and reproducibility standard deviations were calculated, applying Algorithms A and S as described in the ISO 5725-5 standard [15].

2.4.5. Acceptability of the Results Used for Determination of the Assigned Value

Some results were excluded from statistical analysis for the reasons below:

- The second replicate quantification was not reported, which did not allow the precision statistics to be estimated. For example, the laboratory Lab09 submitted the results for only one trial of Week 2.
- The calculation of the logarithmic reduction in the number of microorganisms was incorrect and unverifiable. Two laboratories provided results for *Escherichia coli*, one laboratory for *Pseudomonas aeruginosa*, three laboratories for *Staphylococcus aureus* and *Aspergillus brasiliensis*, and four laboratories for *Candida albicans*, with calculation errors exceeding the acceptable deviation of 0.5 $\log_{10}$ CFU/g recommended by ISO 11930.
- The non-quantitative results were expressed as the limit of quantification (LoQ), including 10 LoQs reported for *Escherichia coli* and *Staphylococcus aureus* and 27 for *Pseudomonas aeruginosa*.

3. Results

3.1. General Results

A statistical summary of the results is presented in Table 2.

In total, 37 laboratories in America, Asia, and Europe participated in this PT to determine the logarithmic reduction in the counts of *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Candida albicans*, and *Aspergillus brasiliensis*. All laboratories applied the same method, which was in agreement with the ISO 11930 standard. Each laboratory determined the number of surviving microorganisms and the logarithmic microbial reduction values in duplicate at an interval of 7 days. As two trials on identical samples were planned at a one-week interval, the results of the first trial were those of the Week 1 test, and those of the second trial belonged to the Week 2 test.

Table 2. Statistical summary.

Test Microorganisms	$p(x_{pt})$ [a]	$x_{pt,1}$ [b]	$\sigma_{pt,1}$ [c]	$s_{r,1}$ [d]	$s_{R,1}$ [e]	$x_{pt,2}$ [f]	$\sigma_{pt,2}$ [g]	$s_{r,2}$ [h]	$s_{R,2}$ [i]	$x_{pt,1} - x_{pt,2}$ [j]	*t*-Statistic	*p*-Value
Escherichia coli	24	2.73	1.04	0.10	1.04	2.56	0.64	0.11	0.64	0.17	0.50	0.618
Pseudomonas aeruginosa	8	3.20	0.70	0.09	0.71	3.15	0.73	0.52	0.82	0.05	0.18	0.860
Staphylococcus aureus	23	3.12	0.40	0.11	0.41	3.17	0.57	0.14	0.58	−0.05	−0.30	0.769
Candida albicans	32	−0.09	0.27	0.05	0.27	−0.11	0.27	0.07	0.27	0.02	0.33	0.739
Aspergillus brasiliensis	33	0.24	0.22	0.06	0.22	0.21	0.24	0.07	0.24	0.03	0.27	0.790

[a] Number of quantitative results reported for both replicate results of Week 1 and Week 2 trials. [b] Assigned value of the Week 1 trial ($\log_{10}$ CFU/g). [c] Standard deviation for proficiency assessment of the Week 1 trial ($\log_{10}$ CFU/g). [d] Repeatability standard deviation of the Week 1 trial ($\log_{10}$ CFU/g). [e] Reproducibility standard deviation of the Week 1 trial ($\log_{10}$ CFU/g). [f] Assigned value of the Week 2 trial ($\log_{10}$ CFU/g). [g] Standard deviation for proficiency assessment of the Week 2 trial ($\log_{10}$ CFU/g). [h] Repeatability standard deviation of the Week 2 trial ($\log_{10}$ CFU/g). [i] Reproducibility standard deviation of the Week 2 trial ($\log_{10}$ CFU/g). [j] Difference in assigned values ($\log_{10}$ CFU/g).

For all test microorganisms, initial counts averaged 6.3 ± 0.1 $\log_{10}$ CFU/g (mean ± SD). The low variability of initial counts demonstrated the sufficient homogeneity of the test samples.

The number of quantitative results varied from 8 to 33 depending on test microorganisms. Only 8 laboratories provided valid quantitative results for *Pseudomonas aeruginosa* in both trials. However, the number of quantitative results (from 23 to 33 results) was sufficiently representative of the results of all participating laboratories for the other microorganisms (*Escherichia coli*, *Staphylococcus aureus*, *Candida albicans*, and *Aspergillus brasiliensis*). Therefore, the assigned values were decided to be estimated from quantitative results reported for both trials. Furthermore, in order to evaluate method precision under repeatability conditions, only the laboratories reporting both replicate results were considered in the calculation of the assigned value.

Only *Pseudomonas aeruginosa* and *Staphylococcus aureus* had microbial reduction values slightly higher than 3.0 $\log_{10}$ CFU/g at day 7 of both trials. Reductions were lower than 3.0 $\log_{10}$ CFU/g for *Escherichia coli* and greater than 0.0 $\log_{10}$ CFU/g for *Aspergillus brasiliensis*, and growth was observed for *Candida albicans* relative to the initial concentration (Table 2). The assigned values demonstrated that the composition of the cosmetic formulation was not effective in its protection against microbial growth according to Evaluation Criterion A of the ISO 11930 standard at day 7, particularly for *Escherichia coli* and *Candida albicans*. However, the conclusion should be confirmed for an analysis of *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Aspergillus brasiliensis* over a period of 28 days, which lies beyond the current framework of this study and should be addressed in future research.

The comparison of replicate results in a PT allowed for evaluation of method precision under repeatability and reproducibility conditions. The robust estimates of repeatability and reproducibility standard deviations were obtained using Algorithms A and S of the ISO 5725-5 standard. In general, the variability of the replicate results measured by the same analyst using the same equipment and applying the same test procedure was low. The repeatability standard deviations were approximately 0.1 $\log_{10}$ CFU/g for most test microorganisms (Table 2). An exception was observed for *Pseudomonas aeruginosa* in Week 2, with the repeatability standard deviation reaching 0.5 $\log_{10}$ CFU/g. The variability of the results was mainly attributed to the reproducibility of results determined by different analysts under varying test conditions, which explained the variation of reproducibility standard deviations in the range of 0.2–1.0 $\log_{10}$ CFU/g.

In order to evaluate the variability of the results over a long-term interval, the whole trial was carried out on identical samples by the same analyst using the same test equipment and analytical procedure one week after the first trial. Welch's *t*-statistics [18] varied from

−0.30 to 0.50, with corresponding *p*-values ranging from 0.618 to 0.860, indicating the insignificant difference in the assigned values of the two trials for all test microorganisms (Table 2).

Table 3 provides a summary of laboratory performance in this PT. The results were acceptable for most laboratories participating in the calculation of the assigned values, with one or two warning results observed for certain microorganisms, and only one unacceptable result of *Aspergillus brasiliensis* was noted in Week 2.

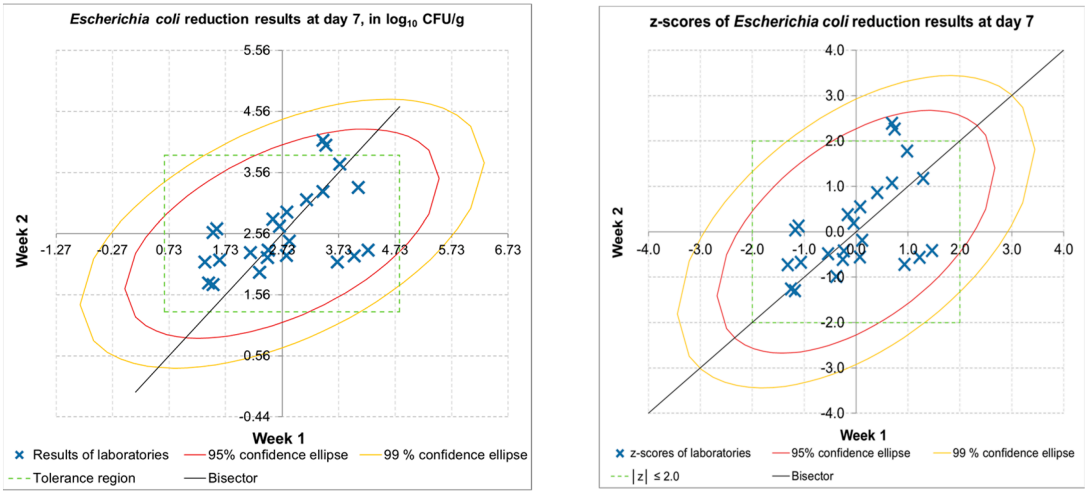
Table 3. A summary of laboratory performance.

Test Microorganisms	$p(x_{pt})$ [a]	$p_{A,1}$ [b]	$p_{W,1}$ [c]	$p_{U,1}$ [d]	$p_{A,2}$ [e]	$p_{W,2}$ [f]	$p_{U,2}$ [g]
Escherichia coli	24	24	0	0	22	2	0
Pseudomonas aeruginosa	8	8	0	0	8	0	0
Staphylococcus aureus	23	23	0	0	23	0	0
Candida albicans	32	32	0	0	31	1	0
Aspergillus brasiliensis	33	33	0	0	32	0	1

[a] Number of quantitative results reported for both replicate results of Week 1 and Week 2 trials. [b] Number of acceptable results in the Week 1 trial ($|z| \leq 2.0$). [c] Number of warning results in the Week 1 trial ($2.0 < |z| < 3.0$). [d] Number of unacceptable results in the Week 1 trial ($|z| \geq 3.0$). [e] Number of acceptable results in the Week 2 trial ($|z| \leq 2.0$). [f] Number of warning results in the Week 2 trial ($2.0 < |z| < 3.0$). [g] Number of unacceptable results in the Week 2 trial ($|z| \geq 3.0$).

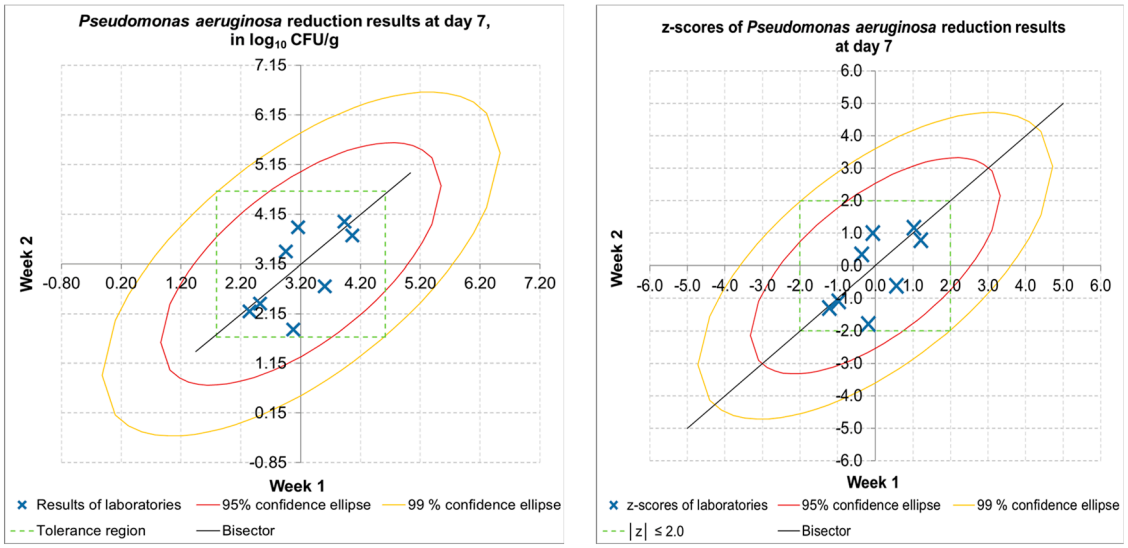
3.2. Individual Laboratory Performance

In addition to statistical analysis, Youden plots were constructed to evaluate the analytical performance of individual laboratories [19–21]. Figures 1a–5a plot the results of the logarithmic reduction in the five test microorganisms at day 7 of Week 1 and Week 2 trials. The abscissa of a point corresponded to a laboratory's logarithmic reduction result in Week 1, and the ordinate was Week 2's result obtained from the analysis of identical samples prepared from the same homogeneous batch of the cosmetic formulation. The intersection of the two coordinate axes exhibited a point whose abscissa was the assigned value of Week 1 and whose ordinate was the assigned value of Week 2. A tolerance region of 2 robust standard deviations from the assigned values, a bisector line, and the 95% and 99% confidence ellipses were also plotted to assess within- and between-laboratory measurement biases.



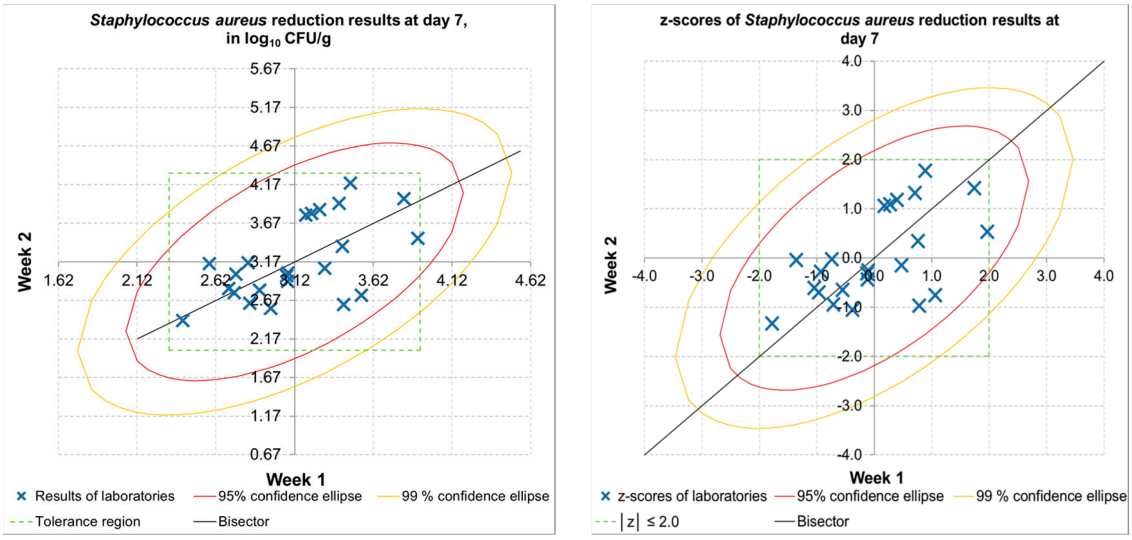
(**a**) (**b**)

Figure 1. Youden plots of (**a**) logarithmic reduction results of *Escherichia coli* obtained at day 7 of Week 1 and Week 2 trials and (**b**) their associated *z*-scores.



(**a**) (**b**)

Figure 2. Youden plots of (**a**) logarithmic reduction results of *Pseudomonas aeruginosa* obtained at day 7 of Week 1 and Week 2 trials and (**b**) their associated *z*-scores.



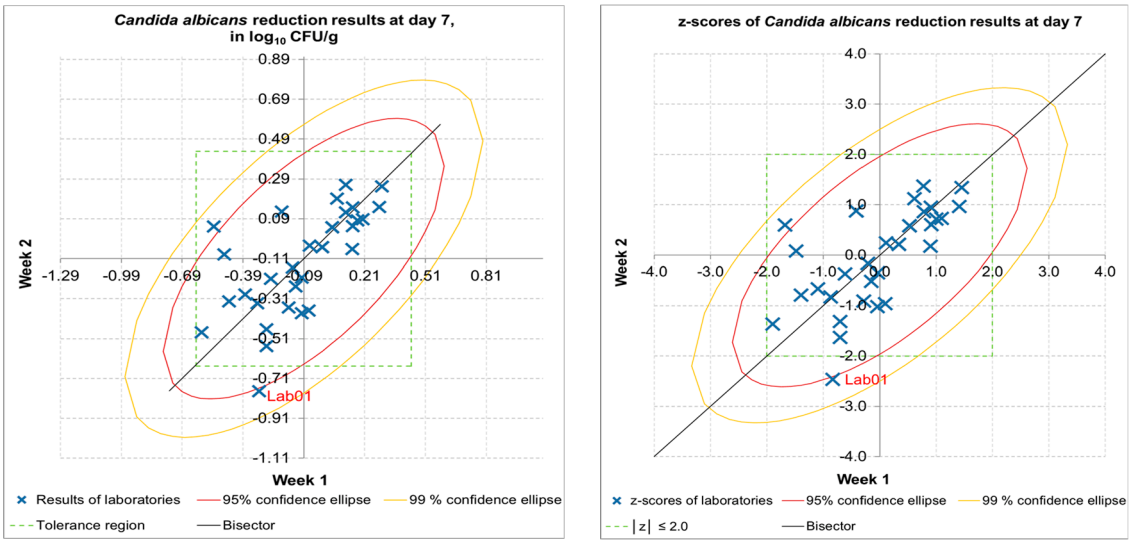
(**a**) (**b**)

Figure 3. Youden plots of (**a**) logarithmic reduction results of *Staphylococcus aureus* obtained at day 7 of Week 1 and Week 2 trials and (**b**) their associated *z*-scores.

In order to facilitate performance comparison, Youden plots of *z*-scores associated with laboratories' results were also plotted. The abscissa of a point on Figures 1b–5b represents the *z*-score associated with Week 1's result, and the ordinate is the *z*-score associated with Week 2's result. The intersection of the coordinate axes was the point of *z*-scores (0.0, 0.0). The region of acceptable results was shaped by a square with edges defined by the *z*-scores of −2.0 and 2.0 on the coordinate axes.

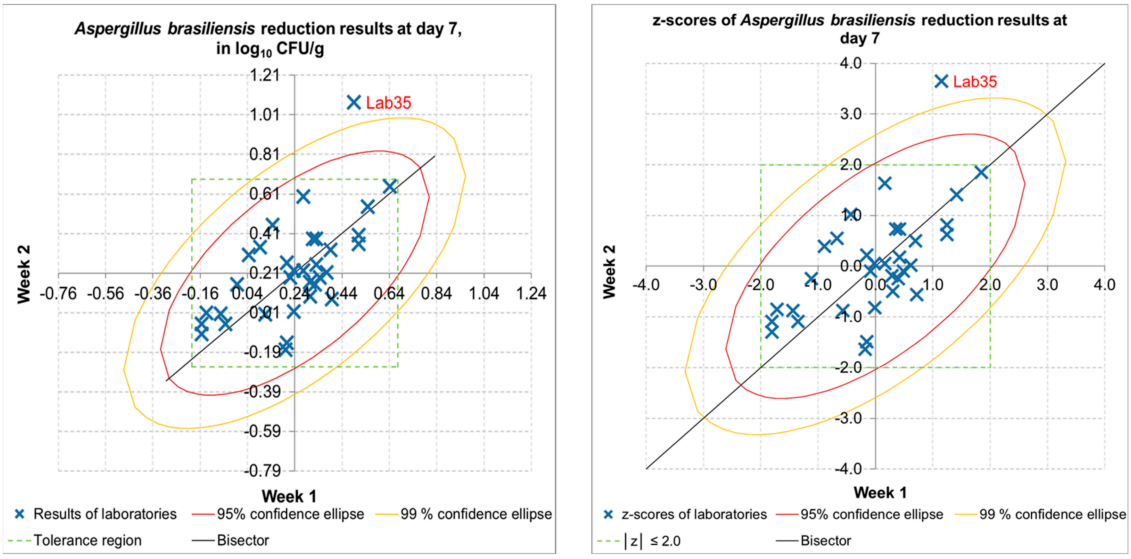
Figures 1–5 show that most laboratories were in the tolerance region and scattered along the bisector line, indicating the acceptable performance of most laboratories and the small variation in Week 1 and Week 2 results. The shape of the 95% and 99% confidence ellipses exhibited a positive correlation between the microbial reduction results obtained on two identical samples at a 7-day interval. As shown in the graphs of *Escherichia coli* and *Candida albicans*, several laboratories fell outside the tolerance and $|z| \leq 2.0$ regions, but they were inside the 95% and 99% confidence ellipses, and they were remote from

the assigned values in Week 2 only (Figures 1 and 4). It was likely that these deviations exhibited random biases. Similarly, the laboratory Lab35 reported a higher result than the results of the other laboratories in Week 2 and obtained an unacceptable z-score of 3.6, yet the laboratory's result in Week 1 was acceptable, which did not imply a systematic bias (Figure 5). However, an internal examination of the cause of this deviation was necessary.



(**a**) (**b**)

Figure 4. Youden plots of (**a**) logarithmic reduction results of *Candida albicans* obtained at day 7 of Week 1 and Week 2 trials and (**b**) their associated *z*-scores.



(**a**) (**b**)

Figure 5. Youden plots of (**a**) logarithmic reduction results of *Aspergillus brasiliensis* obtained at day 7 of Week 1 and Week 2 trials and (**b**) their associated *z*-scores.

4. Discussion

This study was one of the few proficiency tests in antimicrobial efficacy testing for cosmetic products, bringing together a large number of participating laboratories from different countries. Although replicate analysis is not required in ISO 11930, the determination of microbial reduction analyzed in duplicate under both unchanged and changed test conditions allowed for evaluation of both repeatability and reproducibility of the analytical method. In general, the observed repeatability standard deviations of microbial reduction

results were small and comparable across all test microorganisms, demonstrating satisfactory analytical precision. However, reproducibility standard deviations, ranging from 0.2 to 1.0 $\log_{10}$ CFU/g, reflected considerable variability between laboratories.

Given the inherent uncertainty associated with the enumeration of viable microorganisms, determining the "true" concentration of a test microorganism in a cosmetic sample is challenging [22]. Among the five test microorganisms of this study, *Escherichia coli* and *Pseudomonas aeruginosa* presented specific analytical challenges. This could be seen from the high number of results expressed as a limit of quantification for the latter. In addition, the high reproducibility standard deviations (1.04 and 0.64 $\log_{10}$ CFU/g) also demonstrated the difficulties of certain laboratories in the quantitative analysis of these test microorganisms. The variability in results across participating laboratories may be attributable to variations in media and strain suppliers, sampling, equipment and procedures used for sample weighing and homogenization, storage conditions, diluents, dilution protocols, incubation durations, and analyst experience.

The results of this study also indicated that the proficiency test developed for antimicrobial efficacy testing for cosmetic products provided an effective tool for laboratories to evaluate and monitor the quality of their analytical performance through comparison with other participating laboratories. In addition to statistical analysis in accordance with the ISO 13528, ISO 1725-5, and ISO 11930 standards, Youden plots proved their effectiveness in detecting random and systematic measurement biases. For example, the laboratory Lab35 was identified as an outlier in Figure 5. The deviations observed for this laboratory may arise from other internal operating practices during the implementation of the method for determining *Aspergillus brasiliensis* reduction values, and they are laboratory-specific sources of variation. An internal investigation of the causes of these large deviations is necessary for this laboratory, and corrective actions may be taken to assure the reliability of analytical results. Although the evaluation of method precision in this study did not reflect routine laboratory practices, integrating it into a proficiency test highlighted that proficiency testing served not only to assess analytical performance and method reproducibility among different laboratories but also to support the continuous enhancement of quality control procedures in cosmetic microbiology laboratories.

The findings and the design of this study are specific to a cosmetic product with a relatively low level of antimicrobial protection. At day 7, the test cosmetic formulation was considered to be protected against *Pseudomonas aeruginosa* and *Staphylococcus aureus* (with reductions of 3.1–3.2 $\log_{10}$ CFU/g) and *Aspergillus brasiliensis* (with a reduction of 0.2 $\log_{10}$ CFU/g). However, the product was not effectively protected against *Escherichia coli* and *Candida albicans*. To evaluate and confirm laboratory analytical performance, proficiency testing should be extended to a broader range of cosmetic products with varying levels of antimicrobial efficacy over the full 28-day test period in future work.

5. Conclusions

Despite difficulties encountered by some laboratories, the PT demonstrated the overall acceptable performance of the laboratories participating in the determination of the logarithmic reduction in the test microorganisms in a cosmetic formulation potentially contaminated under normal storage and usage conditions. Although the participating laboratories used different broths and agar media and incubated the test microorganisms for varying durations, the results of the laboratories implied the robustness of the protocol for this proficiency testing scheme.

In addition to the main objective of assessing the proficiency of a laboratory, experimental designs of an interlaboratory proficiency test allow the monitoring of various factors that can influence the reproducibility of an analytical method. The reliability of the method

should be confirmed in supplementary trials on various types of cosmetic products with different levels of antimicrobial efficacy, and analysis should be carried out over a long period of 28 days as recommended by the ISO 11930 standard. Future studies should also consider quantifying different sources of enumeration uncertainty (for example, media, equipment, diluents, incubation temperature, incubation time, contamination levels, and antimicrobial protection levels) to achieve a more comprehensive understanding of the uncertainty associated with determining microbial reduction values for the five test microorganisms in a cosmetic product.

Furthermore, regular participation in proficiency tests provides an effective tool for laboratories to control their analytical quality by comparing with other laboratories, promptly identify the anomalies in their routine analytical procedures, and enhance the safety of cosmetic products launched on the market as well as the confidence of consumers and public institutions.

Author Contributions: Conceptualization, A.T., M.F.E., A.B. and S.B.; methodology, A.T., M.F.E., A.B., S.B. and T.T.N.; formal analysis and investigation, S.B. and T.T.N.; writing—original draft preparation, T.T.N. and S.B.; writing—review and editing, M.F.E., A.B. and A.T.; visualization, T.T.N. and A.T.; project administration, S.B.; supervision, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available upon request.

Acknowledgments: BIPEA acknowledges all the laboratories participating in this study.

Conflicts of Interest: S.B., T.T.N., A.T. and A.B. are affiliated with BIPEA (Paris, France), while M.F.E. is affiliated with L'Oréal (Chevilly-Larue, France). The interlaboratory study was conducted by BIPEA in collaboration with L'Oréal. BIPEA was responsible for sample preparation and distribution, as well as data processing and analysis. BIPEA confirms that the results were evaluated impartially. All authors declare that they have no conflicts of interest.

References

1. Herrera, A.G. *Public Health Microbiology*; Spencer, J.F.T., Ragout de Spencer, A.L., Eds.; Humana Press: Totowa, NJ, USA, 2004; Volume 268, pp. 293–295.
2. Alshehrei, F.M. Preservation of cosmetics against microbial contamination. *Mathews J. Dermatol.* **2025**, *9*, 28. [CrossRef]
3. Jairoun, A.A.; Al-Hemyari, S.S.; Shahwan, M.; Zyoud, S.H. An investigation into incidences of microbial contamination in cosmeceuticals in the UAE: Imbalances between preservation and microbial contamination. *Cosmetics* **2020**, *7*, 92. [CrossRef]
4. Bashir, A.; Lambert, P. Microbiological study of used cosmetic products: Highlighting possible impact on consumer health. *J. Appl. Microbiol.* **2020**, *128*, 598–605. [CrossRef] [PubMed]
5. Kim, H.W.; Seok, Y.S.; Cho, T.J.; Rhee, M.S. Evidence from the national pilot project for public health. *Sci. Rep.* **2020**, *10*, 1561. [CrossRef] [PubMed]
6. Duarte da Silva, J.; Silva, F.A.M.; Rodrigues, C.F. Microbial contamination in cosmetic products. *Cosmetics* **2025**, *12*, 198. [CrossRef]
7. *ISO 29621*; Cosmetics—Microbiology—Guidelines for the Risk Assessment and Identification of Microbiologically Low-Risk Products. International Standardization Organization: Geneva, Switzerland, 2017.
8. Halla, N.; Fernandes, I.P.; Heleno, S.A.; Costa, P.; Boucherit-Otmani, Z.; Boucherit, K.; Rodrigues, A.E.; Ferreira, I.C.F.R.; Barreiro, M.F. Cosmetics preservation: A review on present strategies. *Molecules* **2018**, *23*, 1571. [CrossRef] [PubMed]
9. Poddebniak, P.; Kalinowska-Lis, U. A survey of preservatives used in cosmetics products. *Appl. Sci.* **2024**, *14*, 1581. [CrossRef]
10. *ISO 11930*; Cosmetics—Microbiology—Evaluation of the Antimicrobial Protection of a Cosmetic Product. International Standardization Organization: Geneva, Switzerland, 2019.
11. *ISO/IEC 17025*; General Requirements for the Competence of Testing and Calibration Laboratories. International Standardization Organization: Geneva, Switzerland, 2017.

12. *ISO/IEC 17043*; Conformity Assessment—General Requirements for the Competence of Proficiency Testing Providers. International Standardization Organization: Geneva, Switzerland, 2023.
13. *ISO 13528*; Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparison. International Standardization Organization: Geneva, Switzerland, 2022.
14. *ISO 5725-2*; Accuracy (Trueness and Precision) of Measurement Methods and Results—Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method. International Standardization Organization: Geneva, Switzerland, 2025.
15. *ISO 5725-5*; Accuracy (Trueness and Precision) of Measurement Methods and Results—Part 5: Alternative Methods for the Determination of the Precision of a Standard Measurement Method. International Standardization Organization: Geneva, Switzerland, 2025.
16. *ISO 22117*; Microbiology of the Food Chain—Specific Requirements and Guidance for Proficiency Testing by Interlaboratory Comparison. International Standardization Organization: Geneva, Switzerland, 2019.
17. Jarvis, B. *Statistical Aspects of the Microbiological Examination of Foods*; Academic Press: Amsterdam, The Netherlands, 2016.
18. Welch, B.L. The generalization of "Student's" problem when several different population variances are involved. *Biometrika* **1947**, *34*, 28–35. [CrossRef]
19. Youden, W.J. Graphical diagnosis of interlaboratory test results. *Ind. Qual. Control* **1959**, *15*, 133–137.
20. Mandel, J.; Lashof, T.W. Interpretation and generalization of Youden's two-sample diagram. *J. Qual. Technol.* **1974**, *6*, 22–36. [CrossRef]
21. Van Nuland, Y. ISO 9002 and the circle technique. *Qual. Eng.* **1992**, *5*, 269–291. [CrossRef]
22. Noblett, T. Establishing the standard deviation for proficiency assessment (σ) in microbiology PT/EQA schemes. *Accredit. Qual. Assur.* **2012**, *17*, 383–388. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.