Received: 8 April 2024 | Accepted: 3 July 2024

DOI: 10.1111/ics.13007

International Journal of Cosmetic Science
SOCIÉTÉ FRANÇAISE DE COSMÉTOLOGIE

RESEARCH COMMUNICATION

Overview of proficiency testing results for the in vivo determination of sun protection factor

Djigo-Isett Zago[1] | **Souad Ben Bari**[1] | **Anne Tirard**[1] | **Sébastien Miksa**[2] | **Pascale Renoux**[3] | **Emmanuel Questel**[4]

[1]BIPEA, Paris, France
[2]WENEOS, Creil, France
[3]L'OREAL, Chevilly-Larue, France
[4]Pierre-Fabre, Paris, France

Correspondence
Djigo-Isett Zago, BIPEA, 189 rue d'Aubervilliers, Paris 75018, France.
Email: dzago@bipea.org

Abstract

A sunscreen product is allowed to be marketed if a protection is provided against ultraviolets (UV) including UVA rays and UVB rays expressed by the sun protection factor (SPF). UVB is radiation that is in the region of the ultraviolet spectrum which extends from about 290 to 320 nm in wavelength and that is primarily responsible for sunburn, ageing of the skin, and the development of skin cancer. Thus, since April 2009, the Bureau Interprofessionnel d'Etudes Analytiques (BIPEA) set up a proficiency testing scheme (PTS) for the determination of SPF in vivo of sunscreen products according to ISO 24444 standard [Cosmetics - Sun protection test methods - *in vivo* determination of the sun protection factor (SPF)] to evaluate the analytical performances of laboratories on these analyses. This PTS gathers twenty-six laboratories around the world with one trial a year. For each test, the statistical treatment of the data is performed according to ISO 13528 standard [Statistical methods for use in proficiency testing by interlaboratory comparison]. The assigned and tolerance values are calculated from the participants' data and the performances of the laboratories are evaluated individually and collectively according to ISO 17043 standard [Conformity assessment - General requirements for proficiency testing]. This paper presents the design of the PT program, its development, and an attentive analysis of laboratories results, which highlight the global performances obtained by laboratories on this type of analysis. The evaluation of the results shows, in fact, a relatively constant dispersion of data since the implementation of the PT program (variability between 10% and 50%).

KEYWORDS

proficiency test scheme, safety testing, skin barrier, statistics, sun protection factor, sunscreen products

Résumé

Un produit solaire peut être commercialisé s'il offre une protection contre les ultraviolets (UV), y compris les rayons UVA et UVB, exprimée par le facteur de protection solaire (FPS). Les UVB sont des rayonnements qui se situent dans la région du spectre ultraviolet dont la longueur d'onde s'étend d'environ 290 à 320 nm et qui sont principalement responsables des coups de soleil, du vieillissement de la

peau et du développement du cancer de la peau. Ainsi, depuis avril 2009, le Bureau Interprofessionnel d'Etudes Analytiques (BIPEA) a mis en place un système d'essais d'aptitude pour la détermination du FPS in vivo des produits de protection solaire selon la norme ISO 24444 [Cosmétiques — Méthodes d'essai de protection solaire — Détermination in vivo du facteur de protection solaire (FPS)] afin d'évaluer les performances analytiques des laboratoires sur ces analyses. Ce programme d'essais d'aptitude regroupe vingt-six laboratoires dans le monde à raison d'un essai par an. Pour chaque essai, le traitement statistique des données est effectué selon la norme ISO 13528 [Méthodes statistiques utilisées dans les essais d'aptitude par comparaison interlaboratoires]. Les valeurs assignées et les valeurs de tolérance sont calculées à partir des données des participants et les performances des laboratoires sont évaluées individuellement et collectivement conformément à la norme ISO 17043 [Évaluation de la conformité — Exigences générales concernant la compétence des organisateurs d'essais d'aptitude]. Cet article présente la conception du programme d'essais d'aptitude, son développement, et une analyse attentive des résultats des laboratoires, qui mettent en évidence les performances globales obtenues par les laboratoires sur ce type d'analyse. L'évaluation des résultats montre en effet une dispersion des données relativement constante depuis la mise en place du programme (variabilité entre 10% et 50%).

INTRODUCTION

Excessive exposure to the sun is a major cause that augment the skin cancer, erythema, oedema, abnormal pigmentation, and finally suppress the immune system [1, 2]. To protect human skin from UVB rays, it is crucial to determine the sun protection factor (SPF) level. For that, the SPF is calculated by dividing the minimal erythmal dose of a protected skin (MEDp) by the minimal erythmal dose of unprotected skin (MEDnp). The more the SPF is higher, the more the skin is protected from these rays.

Furthermore, a sunscreen product could be chosen depending on several factors as:

- The type of skin for the SPF level (a low SPF level is more suitable for darker skin type),
- The area of skin to protect.
- The homogeneity of sunscreen spread.
- The astronomical situation including the time of the day (between 10:00 AM and 02:00 PM. The UV rays are the most harmful for human skin) and the season. The geographical position in terms of latitude and altitude.
- The meteorological environment through several parameters such as the ozone layer, cloudiness, and pollution.
- The formula (texture, fragrance, etc.), which must appeal and therefore lead to more generous and more frequent application.

The huge variety of offered products makes it easy for customers to find a product that matches their preferences while offering sufficient protection.

The purpose of this work is to describe the design of the proficiency testing programme for the SPF in vivo determination and to present results of ten years of rounds.

MATERIALS AND METHODS

Setting up

A proficiency test (PT) programme is an external quality control to: (i) evaluate the analytical performances of a laboratory according to ISO 17043 standard [3] (the present aim in this article); (ii) validate a reference method (i.e., alternative or internal methods); and (iii) characterize a material (such as for reference standard limits).

Whatever the objective, a proficiency test entails the analysis by different laboratories of the same analytical parameters on identical samples. The setting up of a proficiency test can be schematized by three main steps: preparation of the samples, analyses by the laboratories, and statistical treatment of the data, with the estimation of an assigned value and evaluation of the laboratories performances.

Proficiency testing programme on results for the in vivo determination of sun protection factor was set up in April 2009, and the first trial gathered eight participants. Since then, the participation has continued to increase sign of the growing interest of laboratories in these analyses. The evolution of the geographical distribution worldwide since the first proposed proficiency test is presented in Figure 1. It is important to note that, at the beginning, mainly European laboratories subscribed to this PT, and, over the years, this programme has been a real success worldwide.

Each year, a commission meeting is organized by the BIPEA and gathers the chairperson, the technical group of the committee, and laboratories participating in this PTS. This meeting's objective is to make an overview of results from last trials and choose relevant sunscreen products as samples to be analysed for the next series.

Sample production and shipment

Each batch from a same commercial sunscreen product is homogenized and divided into a series of thirty grams test portions by using a homogenization tun with a whisk. Then a progressive manual filling, which lasts approximately 30 min, ensures the homogeneity of the product between all the samples. Thus, the samples are stored in a cool dark place and sent to participants by hiding any information allowing to recognize the brand name of the product. Participants are informed about an expected level of SPF, which can be different from SPF indicated by producers, as, given the delicate nature of these analyses, a safety margin is taken.

Homogeneity and stability

The homogeneity between samples of produced batches is checked during the step of statistical treatment of the laboratories' data by comparisons between the robust standard deviations of the laboratory results of the studied test from previous ones on similar samples produced according to the same procedure.

The stability of the samples for the duration of the test and even further on is guaranteed by our suppliers if keep in the same storage conditions. Moreover, the examination of the participants' results and the date of analysis are checked during the statistical treatment to confirm there is no variation due to product instability.

Analyses by laboratories

Laboratories receive two samples for the determination of SPF in vivo and must call a panel of volunteers to test samples on them.

A reply form is made available online to allow the laboratories to return their analysis results over a period of two-and-a-half-month period. Participants are asked to analyse samples with expectation of the ISO 24444 standard [4] and to provide further information such as date of the last check of the emission spectrum, the UV source

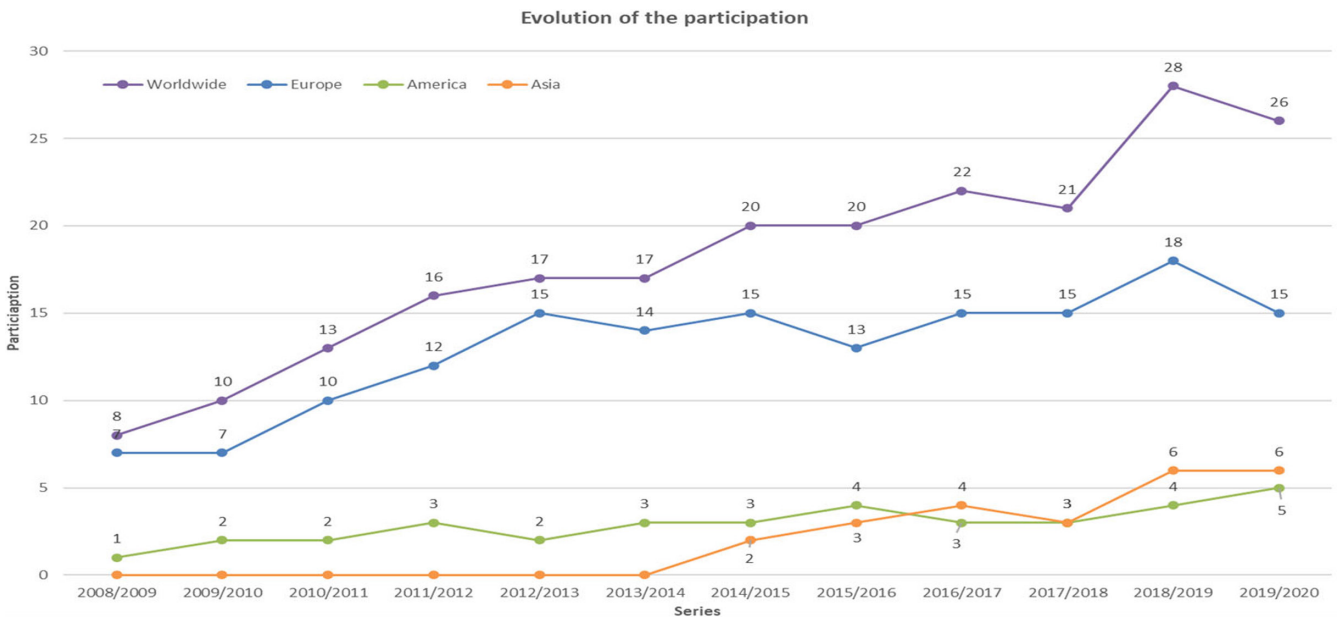

FIGURE 1 Evolution of participation since the 2008/2009 series.

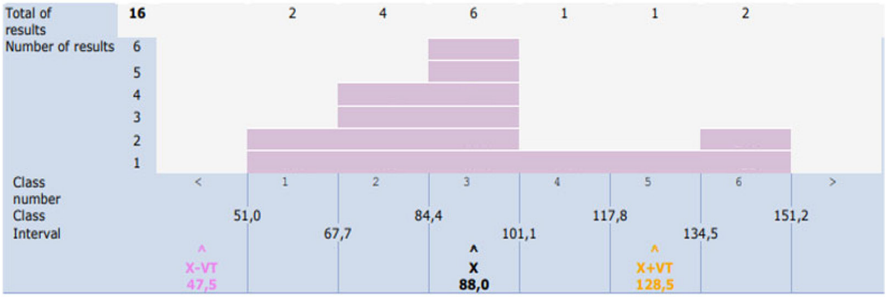

FIGURE 2 Results of daily photoprotection samples represented as a histogram.

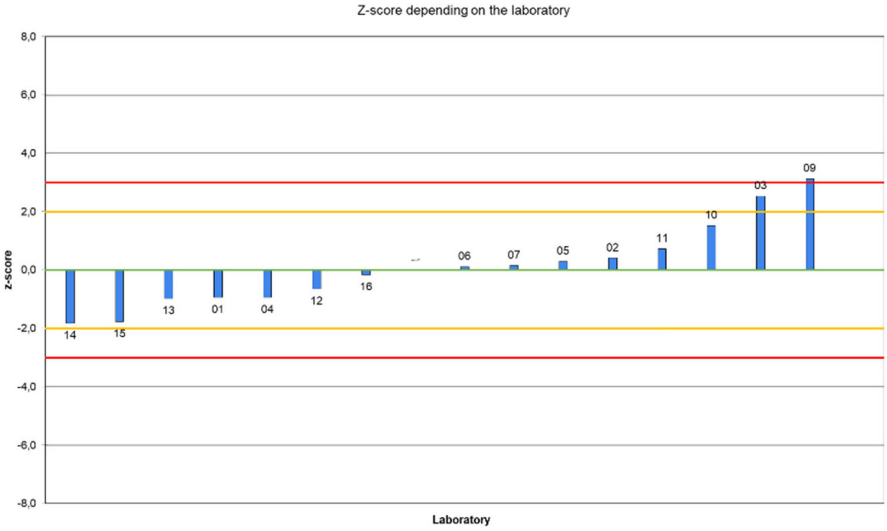

FIGURE 3 Histogram representing *z*-score for all laboratories.

identification, the type of homogenizer used, the reason why, and the number of volunteers excluded.

Statistical treatment

The statistical treatments of the returned results are performed according to the ISO 13528 standard [5]. For each analytical parameter, an assigned value (x_{pt}, conventionally true value, usually named "reference value" in the BIPEA's PTS) is estimated from the robust mean of all the results, except outlier values (e.g., input errors, unit problems, etc.). It is most often calculated by the robust algorithm A from ISO 13528 standard [5]. The proficiency of each laboratory is evaluated thanks to tolerance values (TV). The method used to estimate this value has changed over time: until 2014/2015 series, tolerance values were equal to two times the standard deviation (SD) by robust algorithm A, after 2014/2015 series, and up to now, these values are calculated depending on the assigned value (AV) as follows:

- If AV < 50, the tolerance value is equal to 25% of the assigned value.
- If AV ≥ 50, the tolerance value is equal to 1.5 times the standard deviation by robust algorithm A.

Laboratories' results could be evaluated through *z*-scores (*z*) following the Equation 1.

$$z = \frac{x - x_{pt}}{\frac{VT}{2}} \quad (1)$$

where:

x: measurement result.
x_{pt}: assigned value for proficiency testing.
TV: the tolerance value for proficiency testing.

The *z*-scores can be classified as following:

- If $|z\text{-score}| \leq 2$, the result is acceptable (satisfactory).
- If $2 < |z\text{-score}| < 3$, the result is considered to give a warning signal (questionable).
- If $|z\text{-score}| \geq 3$, the result is considered to give an action signal (unsatisfactory).

Results were published in specific interlaboratory comparison reports distributed to all participants, who could then classify their results and implement some corrective and/or preventive actions if necessary. In this report, each laboratory can be identified through a confidential code to ensure their anonymity.

TABLE 1 Results from all rounds since setting up.

Series	Product	Expected SPF	Assigned value	CV in %	Number of results for the assigned value estimation	Total number of results
2008/2009	Cream	≥25	29.6	18	8	8
	Cream	≥45	58.9	25	8	8
	Cream	≥8	11.2	21	7	7
	Cream	≥15	19.2	15	7	7
2009/2010	Cream	≥8	13.7	28	7	7
	Cream	≥25	N/A	N/A	N/A	6
	Cream	≥25	32.1	14	9	9
	Cream	≥45	55.9	19	9	9
2010/2011	Cream	≥8	11.9	24	10	11
	Cream	≥20	20.1	30	11	11
	Spray	≥25	30.5	20	11	11
	Lotion	≥45	61.8	12	10	11
2011/2012	Oil	≥15	22.5	12	12	13
	Spray	≥25	29.2	20	13	13
	Stick	≥45	68.9	30	9	11
	Cream	≥15	23.0	15	11	12
2012/2013	Cream	≥45	50.8	12	15	15
	Spray	≥45	43.4	15	13	14
	Stick (sensitive areas)	≥25	31.7	15	12	16
	Powder	≥8	7.3	21	13	16
2013/2014	Fundation	≥20	29.9	12	12	14
	Cream (100% mineral)	≥30	49.4	26	13	13
	Milk	≥25	30.2	22	13	13
	Oil	≥25	22.5	12	11	13
2014/2015	Powder	≥30	45.8	20	16	16
	BB cream	≥20	26.6	23	17	17
	Milk	≥45	62.2	22	15	15
	Cream	≥15	22.9	23	16	16
2015/2016	Fundation	≥20	29.5	25	19	19
	Cream	≥40	46.7	12	19	19
	Oil	≥40	43.6	17	15	16
	Spray	≥25	35.0	15	16	16
2016/2017	Children tinted product	≥45	68.2	27	18	19
	Lipstick	≥10	18.9	25	15	18
	Daily photoprotection (moisturizing cream)	≥45	88.0	31	16	17
	Gel for sensitive skin	≥45	100.2	43	15	15
2017/2018	Lotion	≥45	57.8	13	17	17
	Cream	≥15	23.5	23	17	17
	Lotion (with finger cot)	≥45	59.0	20	16	16
	Lotion (without finger cot)	≥45	59.6	22	16	16
	BB cream	≥25	26.1	23	15	15
2018/2019	Sun oil	≥45	N/A	N/A	N/A	16
	CC Cream	≥10	20.4	37	18	18
	Sunscreen cream	≥45	56.4	16	25	25
	Sunscreen gel	≥15	16.5	22	23	24
2019/2020	Lipstick	≥45	57.5	26	16	20
	Mist	≥20	13.4	53	11	14
	Powder	≥15	13.6	31	14	14
	Spray	≥20	30.7	40	13	15

Note: N/A, not applicable; for these tests, no assigned value was estimated.

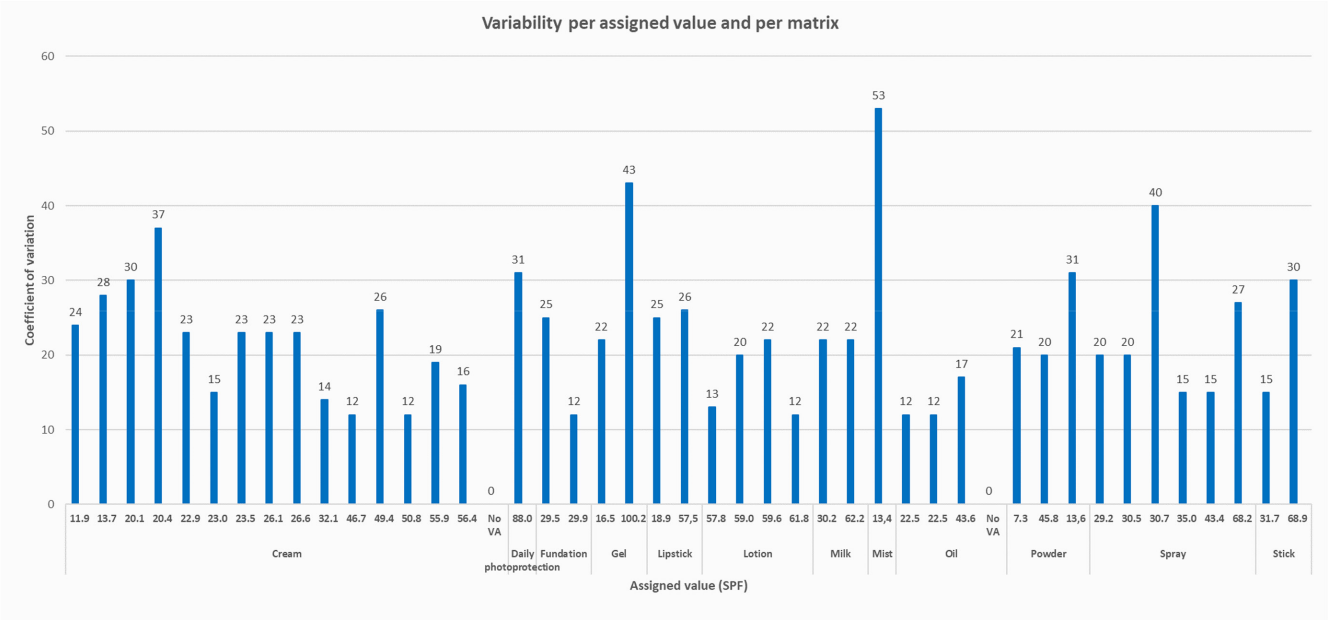

FIGURE 4 Variability per assigned value and per matrix.

Results are also described by graphic form in order to facilitate the reading of laboratory distribution.

Laboratories 'results from the report of the round of March 2017 on a daily photoprotection sample are presented in Figure 2. In this graph, the assigned value ($x=88.0$) and the tolerance range (($x-VT)=47.5$ to ($x+VT)=128.5$) are indicated on the x-axis, whereas the number of results is shown on the y-axis.

Figure 3 presents data obtained in this PT from another perspective and allows each laboratory to know if their result is considered as satisfactory, questionable or unsatisfactory (prior to z-scores). In details:

- Results of laboratories 12 and 10 are considered as satisfactory with a $z \leq |2|$.
- Result of laboratory 03 is questionable with $|2| < z < |3|$.
- Result of laboratory 09 is unsatisfactory with $z \geq |3|$.

Results and discussion

Since the first test, forty-eight tests have already been carried out, and many and various galenics have been analysed by laboratories as sunscreen cream, milk, powder, foundation, oil, stick, mist, spray, gel. Results are summarized in Table 1.

Overall, the results were satisfactory during this decade and the variability (expressed by the percentage of coefficient of variation – CV %) is constant and in the range of 10% and 50%.

Except for the first trials, the assigned values were estimated from at least twelve results, the minimum set for performance evaluation using the robust algorithm A of the ISO 13528 standard. Obtained results also highlight that, for all the products tested, the assigned value is always (except in 2–3 cases) above the expected value, which is generally reassuring for consumers when they buy a product on the market.

However, for three trials, difficulties were faced for participants:

- Cream from the 2008 to 2009 series: no assigned value has been estimated due to the important dispersion among the low number of participants. The difficulties draw the attention of the laboratories once again on the importance of mixing the different phototypes as well as the filtering system of the lamp. The filters that can be used in Europe are filters with a wide spectrum. The difficulties faced on this product have induced the commission to decide to introduce a new parameter on the form. Thus, it was asked to the laboratories to record their percentage relative cumulative erythemal effectiveness (% RCEE) values.
- Sun oil from 2018 to 2019 series: no assigned value could be estimated due to the high dispersion of the results. In fact, two groups of results were observed: one around 30 and the another around 45. The difficulties encountered by the laboratories were mainly due to the type of product and to its composition: an oil whose affinity with the skin could depend on the subjects and lead to migration of the product out of the test area.

TABLE 2 Number of unacceptable results by laboratory according to the number of analysed products.

Laboratory code	Geographical distribution	Number of analysed products	Number of outlier results	Under-estimated unacceptable results	Over-estimated unacceptable results	Total of unacceptable and outlier results
Lab. 22	America	3			1	1
Lab. 14	America	4			1	1
Lab. 21	America	12		1		1
Lab. 4	America	20				0
Lab. 2	America	32	1	2		3
Lab. 9	America	32		1	5	6
Lab. 1	Asia	2				0
Lab. 10	Asia	2			2	2
Lab. 31	Asia	6			2	2
Lab. 35	Asia	6	2		2	3
Lab. 8	Asia	7				0
Lab. 30	Asia	12		3		3
Lab. 26	Asia	18			5	5
Lab. 6	Asia	21			7	7
Lab. 3	Europe	2				0
Lab. 16	Europe	2				0
Lab. 28	Europe	2			1	1
Lab. 13	Europe	4				0
Lab. 33	Europe	4		1	1	2
Lab. 17	Europe	6			1	1
Lab. 11	Europe	8			1	1
Lab. 5	Europe	10			2	2
Lab. 20	Europe	10				0
Lab. 15	Europe	13	1	1	1	3
Lab. 29	Europe	20				0
Lab. 36	Europe	26		2	1	3
Lab. 24	Europe	27			2	2
Lab. 18	Europe	29		2	5	7
Lab. 7	Europe	30			3	3
Lab. 32	Europe	30		2	7	9
Lab. 25	Europe	31		1	5	6
Lab. 23	Europe	33			7	7
Lab. 27	Europe	33	3	1		4
Lab. 12	Europe	34			3	3
Lab. 34	Europe	34			1	1
Lab. 19	Europe	37	1	5	1	7

- Mist from 2019 to 2020 series: the assigned value was estimated for information purpose only. Due to the dispersion of the results, the proficiency assessment was suspended; no tolerance value could be applied. All participants used the same method for this test. The observed dispersion can be linked to the product and not to the method. For this product, a biphasic one, laboratories found difficulties to obtain a homogeneous phase by mixing the product, and some differences of evaporation step were observed.

Figure 4 shows the variability per matrix and SPF value. The variability is satisfactory for most of the tests, even if some CV > 25% can be observed on every kind of

matrices. The matrix with the highest variability is the mist one.

Table 2 shows the number of unacceptable results (outside of the tolerance range) and outlier results (e.g., input errors and unit problems) depending on the number of products analysed by each laboratory.

CONCLUSION

Results' analysis shows that the implementation of this programme is very interesting, not only for laboratories that can now monitor punctually and/or continuously through time the reliability of their results but also to have a global overview of the strengths and weaknesses of these analyses. Currently, this programme proposes two tests per year and gathers more 26 laboratories located in 15 different countries. The number of laboratories registered around the world for this PTS has increased since its launch which shows the growing interest on the analyses of these matrices.

It must be noted that, since the setting up of these PT, a new version of the ISO 24444 standard was published in 2019 [6] with the objective to further improve reproducibility between test sites, to obtain the same SPF value. The main changes compared to the previous edition are as follows:

- The definition of the minimal erythema response (MED) criteria has been revised.
- The choice of eligible test subjects is now based solely on individual typology angle (ITA°) with a requirement for the average ITA° for the test panel to be within the range 41° to 55°, with a minimum of three subjects within two of the three ITA° ranges.
- The ITA° is used to define the range of unprotected MED doses for the provisional or the test day unprotected MED determination (if no provisional MEDu determination is made).
- Three new reference standard sunscreens have been validated and added to the method to validate SPF test panels for products with SPF equal to 25 or higher (P5, P6 and P8).
- New test methods are provided to determine the uniformity of the beam of both large and small beam size solar simulators. A requirement for uniformity greater than or equal to 90% has been added.
- Sunscreen application procedures have been described in greater detail.
- An informative Annex F has been added with photographic examples of erythema responses with guidelines for grading.
- The reporting tables in (Annex G) and the requirements in Clause 11 have been modified to provide more complete information on the results of the testing.

Finally, following the participants' requests, additional developments were carried out to propose PT for UVA in vivo, UVA in vitro and WR analyses. Thanks to these programmes, laboratories can now obtain recognition of their analytical procedures by customers and accreditation bodies according to ISO/IEC 17025 [7].

ACKNOWLEDGEMENTS

The BIPEA acknowledges all laboratories participating in these PTS.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare that are relevant to the content of this article

ORCID

Djigo-Isett Zago https://orcid.org/0009-0006-5683-8728

REFERENCES

1. Narayanan DL, Saladi RN, Fox JL. Review: ultraviolet radiation and skin cancer. Int J Dermatol. 2010;49:978–86.
2. Neale RE, Davis M, Pandeya N, Whiteman DC, Green AC. Basal cell carcinoma on the trunk is associated with excessive sun exposure. J Am Acad Dermatol. 2007;56:380–6.
3. International standard. ISO/IEC 17043:2023 - Conformity assessment - General requirements for proficiency testing.
4. International standard. ISO 24444:2010 - Cosmetics - Sun protection test methods - *in vivo* determination of the sun protection factor (SPF).
5. International standard. ISO 13528:2022 - Statistical methods for use in proficiency testing by interlaboratory comparison.
6. International standard. ISO 24444:2019 - Cosmetics - Sun protection test methods - in vivo determination of the sun protection factor (SPF).
7. International standard. ISO/IEC 17025:2005 – General requirements for the competence of testing and calibration laboratories.

How to cite this article: Zago D-I, Ben Bari S, Tirard A, Miksa S, Renoux P, Questel E. Overview of proficiency testing results for the in vivo determination of sun protection factor. Int J Cosmet Sci. 2024;00:1–8. https://doi.org/10.1111/ics.13007