ALT-SPF

Evaluation and Characterization of Alternative Sun Protection Factor Methods against ISO 24444:2019

I. Introduction

The ALT-SPF project purpose is to evaluate, characterize and compare several alternative SPF test methods to ISO 24444:2019 and to publish the results obtained in peer-reviewed journal(s). The project will take the form of a consortium and therefore a call for expression of interest to join the ALT-SPF consortium (www.alt-spf.com) has been launched on 6 July 2020.

This document aims at sharing with any interested party relevant information regarding the objectives and frame of the ALT-SPF project as well as presenting a suggested approach for the execution of the purpose of the project. This suggested approach will be discussed and validated by the consortium parties.

II. Rationale for the project

The SPF procedure as described in ISO 24444:2019 and the quite comparable procedure in the US as defined by the FDA for SPF measurement are labor-intensive and complex human tests. The UV challenge to the subjects is on a small area but can bear considerable dosages , which is why there has long been a desire for alternative methods. Moreover, the complexity of the procedure and some intrinsic variabilities of this biological test (human subjects, human sunscreen application and human readout) create, for validation purposes, a relatively large random variability, in particular, between different laboratories performing these tests.

Skin cancer incidence rates are still growing around the globe and there are not many silver streams on the horizon that this trend diminishes. The SPF value, measured on human skin, represents a number that is intrinsically linked with the protection against DNA damage in human skin and thus it is connected with the protective ability of sunscreens against skin cancer, although, application of sunscreens should not be the only strategy against the disease. An alternative SPF test method should therefore deliver at least equivalent results as the current SPF *in vivo* in terms of accuracy and there is on top the desire, that the new method shows improvements in precision (repeatability and reproducibility), is easy to adapt by laboratories and cost-effective to perform.

Over the last two to three decades alternative methods such as for example, *in vitro* transmission tests, *in silico* calculation, or non-invasive *in vivo* testing (using negligible UV dose on humans) have emerged and been developed to a stage that makes them promising alternatives.

The desire to understand the accuracy behavior of the alternative methods demand sophisticated statistical tools to analyze and, in particular, to characterize them for a general public use. These statistical tools aim to deliver key features to rate accuracy and to measure reproducibility, degree of systematic errors against different categories of sunscreens. The accuracy, repeatability and reproducibility should be known to decide on the release of a method measuring protective abilities of sunscreens that are linked to skin cancer prevention.

Comparing alternative SPF data against the gold standard requires a solid set of *in vivo* SPF data and thus represents a significant experimental effort that should not be repeated for each alternative method. This is how, in 2018, the main idea for the ALT-SPF project was born.

III. Objectives and results

The overall aim of the ALT-SPF project is to evaluate, characterize and compare several alternative SPF test methods to ISO 24444:2019 and to publish the results obtained in peer-reviewed journal(s).

Key specific objectives:

- **OB1**: To create a smart test design to cope with a minimum of costs but with representative broadness of applicability in terms of sample category and SPF claims.
- **OB2**: To test the SPF samples according to the revised SPF in vivo method ISO 24444:2019.
- **OB3**: To test the same batch SPF samples using different alternative methods.
- **OB4**: To evaluate and characterize alternative test methods using statistically proven models that are in line with ISO 5725.
- **OB5**: To publish the results of the ALT-SPF project in peer reviewed journal(s). Such publication is expected to be made within half a year after statistical analysis of data. Results may serve norming or standardization bodies as a basis to harmonize and develop standard methods for sun protection.

IV. Suggested approach

Note: The final detailed protocol and statistical model will be subject to discussion and validation by the consortium.

Statistical tools for characterization have recently been created in the frame of preparation of this project and will be described therein after

A. Concepts

Comparing methods requires, first of all, selecting a set of samples that adequality represents the spectrum of product types and claim levels found in the market and for which the methods are applicable. Secondly, for each alternative method included in the study, test results will be obtained on the basis of the selected samples and compared to the test results obtained from the same samples with the gold standard, i.e. the ISO 24444:2019.

The alternative methods may show random variation together with some bias in relation to the gold standard. It is essential to characterize both these aspects of error. It is therefore proposed to apply a statistical model that would make it possible to decompose the total error in its different components. An appropriate experimental design should be implemented, to allow the full power of the statistical model to be assessed. The Consortium will consider applying an orthogonal design and a calculation of the different variance parameters by using a mixed model. Figure 1provides an overview of the ALT-SPF proposed approach. Details of the approach are described thereafter.

B. Examples of alternative methods for testing SPF

From a measurement principle three major approaches have been identified so far for testing an SPF. Note that some approaches could contain several specific methods utilizing the same physical principle. There may be other methods, that also deliver SPF data, that the consortium may decide to include in the project.

1. *In vitro* transmission

The sunscreen product is applied on UV transparent templates, e.g. roughed PMMA, that serves a skin mimic. The transmission is recorded by a spectrophotometer. Together with the erythema action spectrum an SPF factor is calculated. Photostability is factored in by specific radiation challenges and remeasurement of transmission spectrum.

2. *In silico*

In silico procedures calculate the transmission by individual contributions of each UV filters and also in relation to the erythema action spectrum and an artificial light source spectrum. The key component is that *in silico* procedures model the sunscreen film roughness on human skin digitally by calibration with experimental *in vivo* data and factor this function in to modulate the transmission spectrum composed from individual UV filter components accordingly to yield an *in silico* SPF that aims to match *in vivo* SPF data.

3. Hybrid diffuse reflectance spectroscopy (HDRS)

HDRS is based on the non-invasive diffuse reflectance spectroscopy (DRS) involving human subjects, however, the radiation UV burden is negligible. UVA protection (320-400nm) is directly assessed from the reflectance spectrum *in vivo*. Since the returning signal is not sufficient in UVB (290-320nm), the SPF is calculated by extrapolating the UVA curve by using *in vitro* measurement, which makes it the hybrid method.

C. Proposed experimental design of the study and statistical analysis of the data

It is not possible to adequately characterize and assess the performance of an alternative method against the gold standard if the observed error is not partitioned into different components, such as laboratory bias, repeatability error and variation between products or matrices. Indeed, even in the ideal case that the alternative method is perfect, meaning both that it has no bias in relation to the gold standard and that its precision is very high, deviations between the two methods may nonetheless be observed, due to different error sources such as product-specific laboratory bias and random repeatability error affecting the results of the gold standard.

Breaking down observed differences into their different constituents will allow the correct assignation of errors to their respective underlying causes. For instance, if overall bias is negligible, but a laboratory bias affects the gold standard's results, individual differences can be correctly interpreted as caused by the gold standard rather than the alternative method. Consider another example: for an alternative method that is built on UV spectra of the components of a sunscreen, showing neither a repeatability error nor a laboratory bias. However, the preliminary determination of absorption values may have been subject to errors. Take the case that for a given formulation ingredient A, the absorption value made available in the method's reference data base has a negative bias. Then, in all products containing a certain amount of formulation, ingredient A will be affected by this negative bias. This example shows that the distinction between different sources of error is important in correctly characterizing method performance.

To put it succinctly: in many cases, and in particular if both the gold standard and the alternative method display considerable variation, a correct characterization and assessment of method performance is not possible without such a decomposition of observed error. Only this level of understanding of a method's behavior will allow an adequate and transparent prediction of how it will perform when made available to test institutes.

Accordingly, it is proposed to implement an experimental design which allows an adequate characterization of the effects associated with laboratories, products and repeatability error. More specifically, the following setup is proposed:

- Sample dimension 1: Three different product types:
  - all emulsion SPF product types including W/O, W/Si and O/W with a maximum of 5 % solid pigments.
  - one-phase products like sunscreen oils, sprays and aerosols.
  - high-solid content formulas, e.g. all mineral sunscreens or formulas with more than 5 % particulate UV filters, ideally more than 10 %.

The product types of dimension 1 are created by looking at what is actually tested when assessing the SPF *in vivo*. It is in a lot of cases not the product that is contained in the dispenser anymore. Evaporation of mainly water, but also alcohol leads in the *in vivo* experiment to changes in the overall composition and the creation of the final sunscreen film on skin. Emulsions lose water very fast when applied to the skin. Practically all water evaporates within 15 minutes of application and a continuous hydrophobic film containing the sunscreen actives is formed, no matter if the emulsion was O/W or W/O. The different emulsion categories were therefore bundled in one group. One-phase products form a separate group as the residual viscosity of those products after application might be very low and differs from emulsions. Since pigments can arrange themselves differently on skin and on templates, high-pigment formulas constitute yet another challenge and are represented in a third product type. The three product types cover more than 90 % of all sunscreen formulas in the market and most skin care products offering sun protection.

- Sample Dimension 2: For each product type, it is suggested that three different SPF levels are considered
  - SPF 15
  - SPF 30
  - SPF 50+ (minimum SPF 60)

These three SPF categories follow a geometric series and represent a large majority of the products in the market. It will likely be possible to use the data obtained from these three SPF levels products to characterize method performance at SPF values below 15 or lying between the levels included in the design.

For statistical reasons, a minimum of four products per combination of product type (Dimension 1) and SPF level (Dimension 2) is required. Accordingly, a total of (3 product types x 3 SPF levels x 4 products =) 36 different products is proposed. Samples will be selected mainly from commercial products or standard samples and a decision about selection will be agreed by the consortium. Individual test items will then be taken from the samples, transferred to identical recipients, coded, and sent out to test facilities. Correlated information regarding the commercial identity of the samples and the SPF results will be disclosed neither to the participants, nor to the consortium; nor will it be published later.

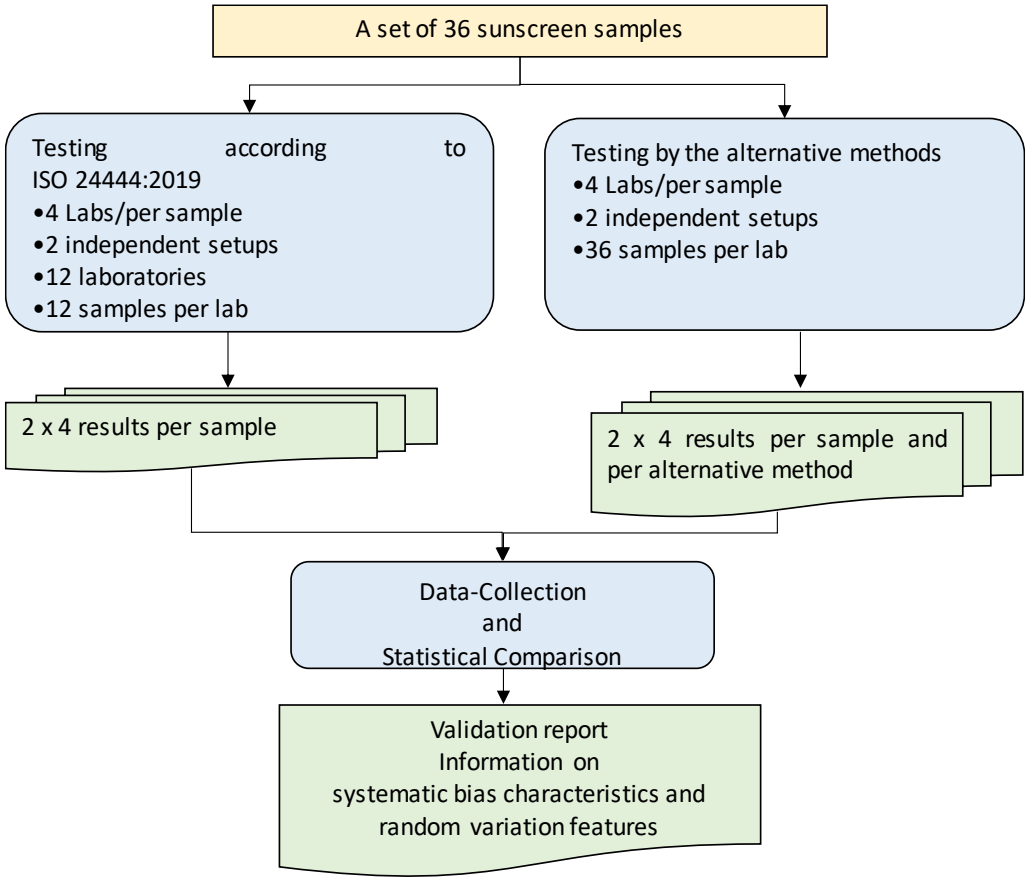


Figure 1: ALT-SPF proposed approach

Since this type of design can result in a considerable workload, great attention will be given to the question of the experimental design's efficiency. In particular, the use of an *orthogonal* design allows a reliable estimation of the different components of variability whilst ensuring economy of resources. More specifically, the orthogonal design allows a reliable estimation of precision components with the possibility to decrease the number of participating institutes to four per method.

The orthogonal design can be implemented as follows. As already explained, for each combination of product type and SPF level, four products are included in the design. For each product, two different settings of three factors will then be implemented, resulting in a total of eight different settings per method. One result is then obtained per setting and method. This is illustrated in Table 1.

Table 1: Suggested design for one combination of product type and SPF level, e.g. for a test institute applying the in vivo method.

	Product 1		Product 2		Product 3		Product 4	
	Setting 1	Setting 2	Setting 3	Setting 4	Setting 5	Setting 6	Setting 7	Setting 8
In vivo test result								

Each setting corresponds to one realization of an orthogonal design involving three different *analytical factors*. For instance, in the case of the *in vivo* method, *Setting 1* will correspond to a particular combination of the three factors *Operator*, *Lamp*, and *Panel of volunteers*.

For each combination of product type and SPF level, this design is implemented in the same way in 4 different laboratories.

Under the proposed design and for the proposed statistical analysis, for each individual alternative method, a minimum of four laboratories is necessary. The laboratories should be able to provide two statistically independent setups in their facilities. For this purpose, 3 relevant factors that influence the result need to be identified and varied to form these two settings, otherwise more laboratories would be needed. For each method, eight test results would therefore be generated for each of the 36 samples.

Going hand-in-hand with the experimental design, the use of a so-called mixed linear model would allow the different bias and precision estimates to be calculated: repeatability error, overall bias, laboratory-specific bias and product bias. In particular, mixed linear models allow a characterization of the variation between laboratories and product bias values to be characterized. Indeed, while both laboratory bias and product bias are interesting as individual positive or negative values affecting a given laboratory or product, it is equally informative to look at the variation across the bias values. In other words, in method evaluation, it is often useful to characterize product or laboratory bias as a standard deviation – just as in the case of repeatability error. The between-laboratory standard deviation can then be used to predict expected variation in bias when SPF is determined for one and the same product in different laboratories, while the between-product standard deviation can be used to predict expected variation in bias when SPF is determined for different products in one and the same laboratory.

V. Benefits of the project

The data generated by the ALT-SPF consortium and the publication of the results may serve as one of the bases in the norming or standardization organizations or other recognized bodies to harmonize and develop standard methods for sun protection that have been proven to be accurate and applicable.

The SPF testing on humans assesses mainly (90 %) damage originating from DNA mutations, that lead to sunburned cells, which trigger by their apoptosis an inflammation reaction that in the end serves as readout. Through this connection, the SPF is strongly linked to the protection against mutagenic activity of UV light. Therefore, it is essential to evaluate and characterize alternative methods thoroughly on their degree of agreement with the *in vivo* SPF. This will ensure that no protection gaps exist because a method overestimates the SPF for a specific category of samples. ALT-SPF will therefore support all current efforts done to prevent skin cancer now with reliable alternative methods. Furthermore, alternative methods will remove the irradiation burden on human subjects from the SPF measurements.

On the other hand, if an alternative method systematically underestimates the SPF, it will force manufactures of sunscreens to unnecessarily add more UV filters to match the *in vivo* SPF claim and would lead to an unnecessary environmental burden.

Accurate and applicable alternatives to the current *in vivo* SPF method will become important tools for screening, claim substantiation and/or in-market control. Through the participation, and later on the publication, test laboratories will have information on how alternative methods operate and their advantages and limits, eventual bias, thus bringing them to the frontline of SPF testing innovation and enabling a choice from a spectrum of methods for a particular type of product.

Alternatives have the potential to be faster, easier and eventually more cost-effectively providing SPF data than the *in vivo* test. Already, during sunscreen development, candidate formulations could be faster selected and development time in total could be shortened, allowing sunscreen manufacturers to reduce the time to market.

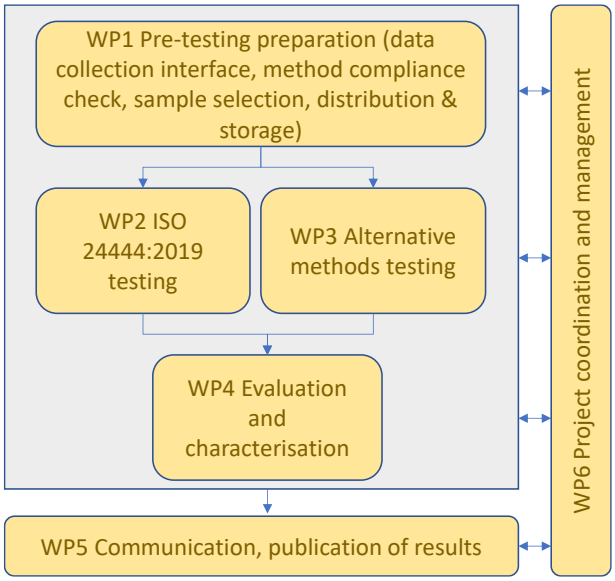
VI. Work plan structure and timing

The project work plan is divided into three main phases, which are slightly overlapping:

1. Pre-testing preparation (WP1; 3 months)
2. Testing of SPF using gold standard and alternative methods (WP2 and 3; 7 months)
3. Statistical evaluation and characterization (WP5; 3 months)

Throughout the project, dissemination and communication will be made to main stakeholders in the WP5, which will also prepare the final publication in a peer-reviewed journal at the latest six months after statistical evaluation.

A specific WP, WP6, will take care of the project coordination and management.



VII. Budget and efforts

The contribution of members can be done in three ways:

1. Sponsoring of project costs (a fixed amount to be calculated based on budget and number of sponsors before project launch)
2. In-kind contribution by carrying out *alternative* testing
3. A mix of 1 and 2.

Please note that the contribution under the conditions set forth in point 2 will give the right to participation to consortium meetings but no voting rights.

A bank account will be held centrally from which payments to service providers, specific consumables, etc. will be made. The detailed functioning of the payment transfers will be setup and established in the Consortium Agreement to be signed by all partners before project launch.

VIII. Further information

For more information about the call, please contact: alt-spf-coordination [at] eurtd.com

For general questions about the organization/management of the project, please contact:

Sara Skogsater, ARTTIC, email: skogsater [at] arttic.eu

For more information related the technical aspects of the project, please contact:

Jürgen Vollhardt, DSM, email: Juergen.Vollhardt [at] dsm.com

Uli Ostenwalder, SPF, email: uli.osterwalder [at] sunprotection.co