Dermatol Ther (Heidelb) (2022) 12:2531–2546
https://doi.org/10.1007/s13555-022-00815-w

ORIGINAL RESEARCH

Evaluation of a Sunscreen Product Compared with Reference Standards P3, P5 and P8 in Outdoor Conditions: a Randomized, Double-Blinded, Intra-individual Study in Healthy Subjects

Corinne Granger · Gitanjali Petkar · Muzzammil Hosenally · Javier Bustos · Carles Trullàs · Thierry Passeron · Jean Krutmann

Received: August 5, 2022 / Accepted: September 7, 2022 / Published online: September 29, 2022

ABSTRACT

Introduction: The shortcomings of standardized sunscreen testing have been discussed in recent years, noting differences between how sunscreens perform in indoor clinical (in vivo) laboratory testing compared with real-life conditions. We previously developed an outdoor clinical method for ranking sunscreens by performance level. We used this method to test the performance of a new broad-spectrum sunscreen against International Organization for Standardization (ISO) reference products P3, P5 and P8.

Methods: Sixty-five healthy volunteers with individual typology angle (ITA) ≥ 28° (light to intermediate skin colour) participated in an outdoor study in Mauritius. Test areas were marked on their backs, which were treated with the different products: one commercially available broad-spectrum sun protection factor (SPF) 50 sunscreen [investigational product (IP)] and the three reference products P3 (SPF 15), P5 (SPF 30) and P8 (SPF 50+) from ISO norm 24444:2019 for SPF testing. The test areas were exposed for 2–3 h, depending on the baseline skin colour. They were also compared with an unprotected positive control area and a non-exposed negative control area. Clinical and colorimetry assessment of erythema and pig-

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13555-022-00815-w.

C. Granger
Stella Polaris Europe, Paris, France

G. Petkar · M. Hosenally
Centre International de Développement Pharmaceutique, Phoenix, Mauritius

M. Hosenally
Department of Economics and Statistics, University of Mauritius, Réduit, Mauritius

J. Bustos · C. Trullàs (✉)
Innovation and Development, ISDIN, Barcelona, Spain
e-mail: carles.trullas@isdin.com

T. Passeron
Department of Dermatology, CHU Nice, University Côte d'Azur, Nice, France

T. Passeron
C3M, INSERM U1065, University Côte d'Azur, Nice, France

J. Krutmann
IUF Leibniz Research Institute for Environmental Medicine, Dusseldorf, Germany

J. Krutmann
Medical Faculty, Heinrich-Heine-University, Dusseldorf, Germany

mentation were performed at 24 h and 8 days, respectively.
Results: Overall, according to this outdoor clinical testing method, the sunscreens' efficacy was ranked in an appropriate order given their established SPF levels, with higher SPFs giving greater protection against erythema and pigmentation. Between the different levels of SPF, the differences were statistically significant, for both clinical and colorimetry assessments. The new broad-spectrum SPF 50 IP performed similarly to the SPF 50+ (P8) reference product. Even the highest SPF products, SPF 50 and SPF 50+, had some instances of photoprotection failure.
Conclusion: These findings confirm the feasibility of this outdoor clinical testing method in ranking sunscreens and provide further evidence, in addition to standardized SPF and UVA protection factor (UVAPF) testing, on how this new broad-spectrum SPF 50 sunscreen performs in extreme outdoor solar exposure: in line with reference product P8 (SPF 50+).
Trial Registration No.: ISRCTN95394014.

Keywords: Extreme; Outdoor conditions; Spectrum; SPF; UV index

Key Summary Points

Why carry out this study?

Critics argue that standard indoor testing of sunscreen efficacy does not include the whole solar spectrum, wavelengths of which can also cause skin damage including erythema and pigmentation. It has been suggested that indoor SPF determination methodologies should be developed to give a more reliable prediction of the sun protection offered in real-life conditions.

This study aimed to determine the efficacy of a new broad-spectrum SPF 50 sunscreen in comparison with reference sunscreen products in outdoor conditions of high UV exposure.

What was learned from the study?

The new broad-spectrum SPF 50 (IP) performed similarly to the SPF 50+ reference product. Differences between adjacent sunscreen levels were statistically significant.

This method seems to align well with SPF categories. The results provide additional evidence on the quality of the sunscreen tested in relation to other reference products and underline the interest of the evaluation of sunscreens in outdoor conditions. None of the products protected fully, so protective measures such as covering with clothes should be encouraged.

INTRODUCTION

The deleterious effects of solar radiation on the skin, in particular ultraviolet (UV) B and A, are well established, and knowledge of the effects of visible light is also expanding [1]. Sunscreens form a key part of the various behaviours that protect against skin damage. To allow consumers to identify levels of protection, sun protection factor (SPF) is used on labelling,

mainly as a measure of UVB protection, and international standardized methods for SPF testing are in force, such as ISO 24444:2019 [2]. There are also standardized methods for UVA testing such as ISO 24442:2011 and ISO 24443:2021 [3, 4]. However, there are some criticisms of the existing methods [5], with evidence to show that SPFs may overestimate the protection offered in natural light [6]. SPF testing has been designed to use only the most relevant wavelengths of light for testing objectives, and therefore differs from real-life exposure as subjects are not exposed to the full spectrum of light [6], yet other wavelengths (namely visible light) have been shown to induce erythema and pigmentation in the skin [7, 8]. In addition, despite following the same standardized methods, inter-laboratory variability in reported SPF values has been demonstrated and discussed [9, 10]. This has been partially attributed to the physical and chemical characteristics of the products and their UV filters [9]. So, existing standardized testing methods are recognized to bear considerable random variability, in particular between different laboratories performing these tests, and must be improved [10].

Our group previously developed an outdoor testing method, carried out in conditions of extreme solar exposure in Mauritius, using clinical and instrumental assessment of erythema and pigmentation to determine the performance of sunscreen products under these conditions [11]. The method was able to rank in order the performance of the products tested on the basis of the clinical and colorimetric assessments, and provided a further level of information to supplement standard SPF and UVAPF testing. In a second study in Singapore in 2021, we modified the method slightly [12] and demonstrated that the method, in particular the clinical scoring, was able to discriminate between the three reference products P3 (SPF 15), P5 (SPF 30) and P8 (SPF 50+).

The aim of the present study was to further fine-tune this method and use it to test a new broad-spectrum SPF 50 sunscreen [investigational product (IP)], to provide evidence of how it actually performs in outdoor conditions of extreme solar exposure compared with established reference products from the ISO norm 24444:2019.

METHODS

Study Design and Population

This was a single-centre, double-blind, randomized, intra-individual clinical study conducted in an outdoor facility in Tamarin, Mauritius (20.3378° S, 57.3751° E) between November 2021 and December 2021 (summer). Seventy-four subjects were enrolled. Inclusion criteria were age 18–55 years, individual typology angle (ITA) $\geq$ 28°, with uniform skin colour throughout the investigational area (ITA $\pm$ 4° difference between any two zones), and absence of dermatological disorders affecting the investigational areas (multiple nevi, freckles, excess hair or uneven skin tones, tattoos, vitiligo or other pigmentary disorders). Individuals with a history of skin cancer, abnormal response to sun, cosmetic allergies or hypersensitivities, or those taking medications likely to interfere with the study outcomes, were excluded.

Products Tested

IP: a sunscreen containing a combination of lipophilic organic sun filters and a hydrophilic organic sun filter, formulated in a new water-rich oil-in-water emulsion. This product is marketed as a broad-spectrum, SPF 50 on the basis of standardized SPF testing with a label SPF of 50 and a UVAPF of 22.8. The reference products were taken from ISO 24444:2019 SPF testing method [2]: P3, which has an SPF of 15 and UVAPF of 2.5; P5, which has an SPF 30 and UVAPF of 13.4; and P8, which has SPF 50+ and UVAPF 27.5. The UVAPF evaluations were measured in an external laboratory according to ISO 24444:2012. Supplementary Table S1 presents the composition of the products.

Sun Exposure and Protection

Test areas were on subjects' backs, with six areas per subject. Each area was treated with a

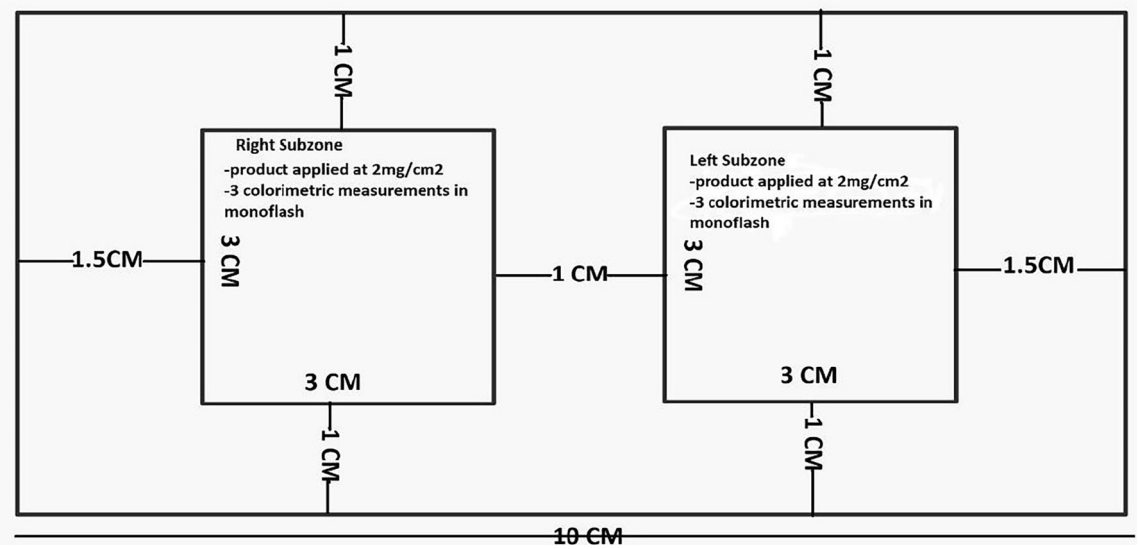


Fig. 1 One assessment area, split into two sub-areas of 3 × 3 cm. Each subject had six assessment areas on their back. In subjects with ITA ≤ 41°, in each test area, one of the sub-areas was exposed for up to 3 h. In the remaining subjects, all treated sub-areas were exposed for 2 h. In all subjects, the untreated exposed area was smaller (1.5 cm^2) and covered after 1 h

Table 1 Clinical scales used for erythema and pigmentation grading

Erythema	
Grade	**Description**
0	No erythema
1	Equivocal reaction, slight, barely perceptible erythema (not clearly defined or does not cover the entire exposed area)
2	Clearly visible erythema with well-defined borders
3	Moderate erythema
4	Severe erythema
5	Very severe erythema with blistering
Pigmentation	
Grade	**Description**
0	No difference with surrounding skin
1	Slight increase (barely visible) in pigmentation
2	Mild increase in pigmentation
3	Marked increase in pigmentation
4	Maximal increase in pigmentation

different sunscreen, according to a randomization list: either the IP (SPF 50), P3, P5 or P8, or no sunscreen (two areas with no sunscreen). Products were applied at 2 mg/cm^2. After application, 15–30 min elapsed before exposure. One of the no-sunscreen areas was covered and not exposed at all, acting as a negative control. The other was exposed for 1 h, acting as a positive control. This exposure was limited to 1 h to avoid excessive burns.

The sunscreen-treated areas were exposed for 2 or 3 h, depending on the subject's baseline ITA, as follows: each treatment area was subdivided into two (Fig. 1), a left and a right. In subjects with ITA ≤ 41° (intermediate colour), one of these sub-areas was covered after 2 h exposure and the other remained uncovered for a further 1 h. In subjects with ITA > 41° (light colour), all exposure stopped after 2 h. The expected cumulative dose of erythemally weighted UVB was 100–250 mJ/cm^2.

If at any time an area developed an immediate erythema score of ≥ 2, it was covered. Subjects lay prone outdoors for the duration of the study. All non-investigational areas were protected with clothing, hats and sunglasses. After exposure, the IP and reference products

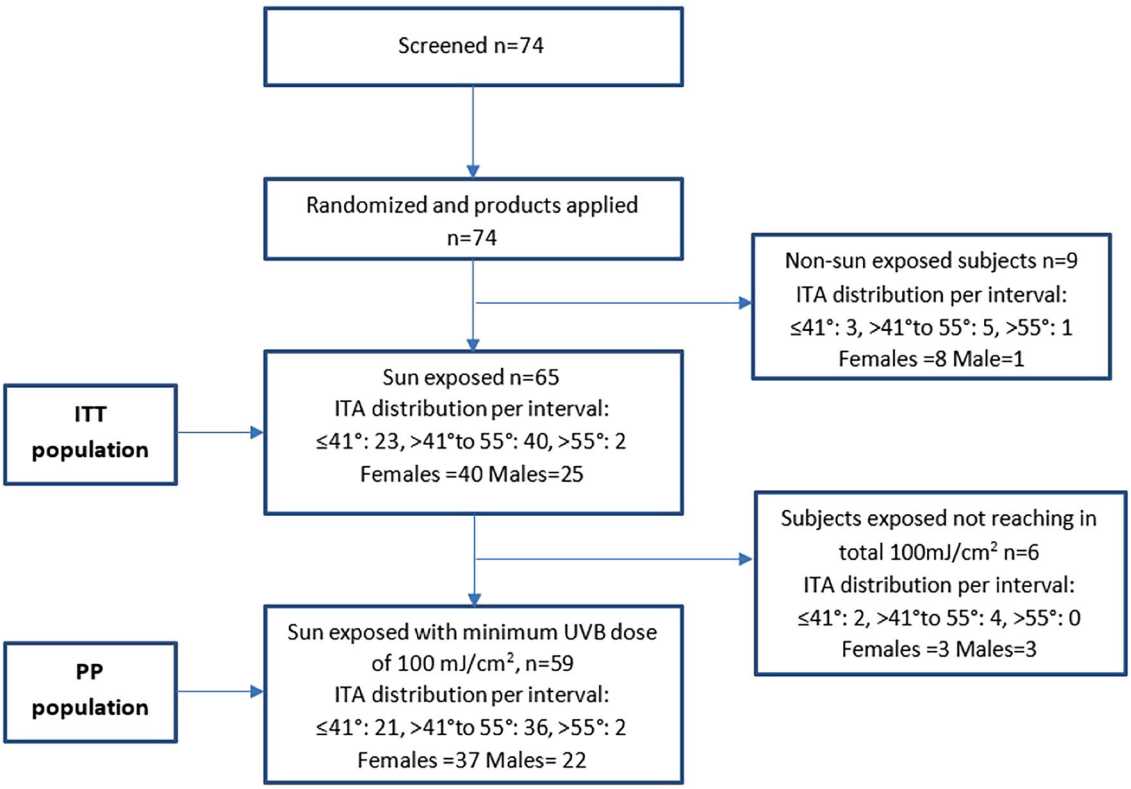


Fig. 2 Study population. ITT population $n = 65$

were removed with water-based wipes. Subjects were instructed to avoid sun exposure and application of any topical products to the study area until the end of the study at day 8.

For each subject, the cumulative UVA and UVB doses were recorded using a radiometer (PMA2100, Solar Light Company Inc, PA, USA) which had UVA and UVB sensors and software to calculate cumulative UVA (J/cm^2) and erythemally weighted UVB (mJ/cm^2). The UV index was calculated from the values of erythemal UV irradiance measured by the radiometer. Irradiance was measured every 10 min during exposure. Temperature and hygrometry were recorded using a calibrated thermo-hygrometer device (TESTO model 174-H SE & Co-KgaA, Lenzkirch, Germany); both were measured in the same area where sun exposure took place.

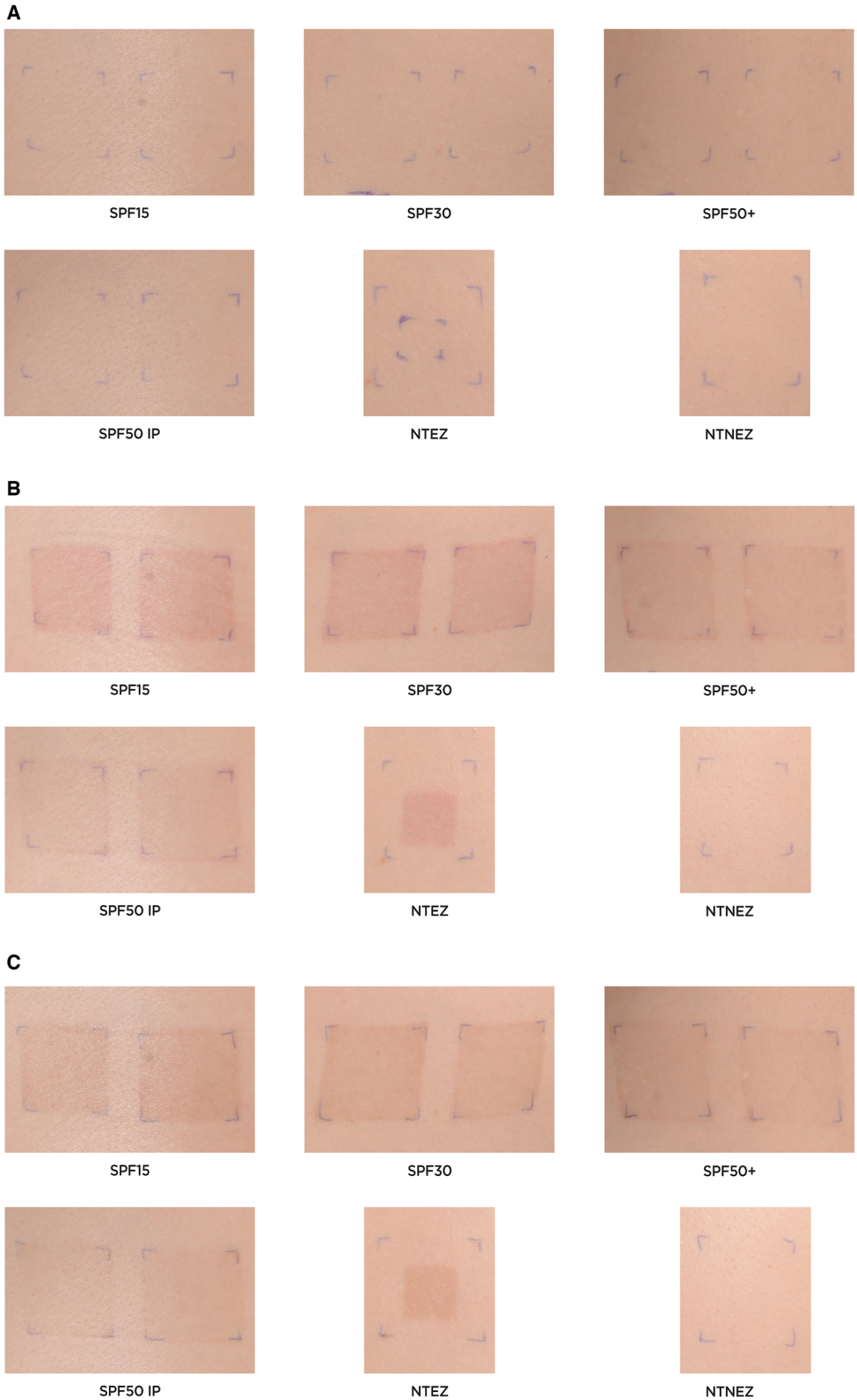
Outcomes

The primary outcome was clinical erythema score (grades 0–5, Table 1) at 24 h after exposure.

Secondary outcomes were delayed pigmentation (grades 0–4, Table 1) [12] and the colorimetry parameters a^* (redness), L^* (lightness) and ITA (overall colour), measured using a Chromameter CR-400 (Konica Minolta, Inc., Tokyo, Japan). Assessment of clinical pigmentation (delayed tanning) was performed at 1 week after exposure (day 8). Colorimetry was performed at both 24 h (20 ± 4 h) and day 8; results are reported for a^* at 24 h, and L^* and ITA at day 8.

Photoprotection failure was defined as a clinical erythema score of 2 or more.

Prior to any assessment, subjects were acclimatized for at least 15 min in a temperature- and hygrometry-controlled room with a temperature of 24 ± 2 °C and hygrometry of 50 ± 10%. Clinical assessments were performed


A
SPF15
SPF30
SPF50+
SPF50 IP
NTEZ
NTNEZ
B
SPF15
SPF30
SPF50+
SPF50 IP
NTEZ
NTNEZ
C
SPF15
SPF30
SPF50+
SPF50 IP
NTEZ
NTNEZ

◀**Fig. 3** Representative photographs of one subject (R039). **A** At baseline; **B** at 24 h; and **C** at day 8. NTEZ, non-treated exposed zone; NTNEZ, non-treated non-exposed zone. Erythema scores at 24 h: SPF 15, grade 3; SPF 30, grade 2; SPF 50+, grade 0; IP, grade 0; NTEZ, grade 3; NTNEZ, grade 0. a* at 24 h, change from baseline: SPF 15, 7.11; SPF 30, 5.29; SPF 50+, 1.46; IP, 1.81; NTEZ, 8.64; NTNEZ, 0.57. Pigmentation scores at day 8: SPF 15, grade 1; SPF 30, grade 1; SPF 50+, grade 0; IP, grade 0; NTEZ, grade 2; NTNEZ, grade 0. L* at day 8, change from baseline: SPF 15, −4.42; SPF 30, −4.42; SPF 50+, −2.14; IP, −1.61; NTEZ, −6.21; NTNEZ, −2.76. ITA at day 8, change from baseline: SPF 15, −8.93; SPF 30, −7.44; SPF 50+, −2.2; IP, −1.62; NTEZ, −12.98; NTNEZ, 0.23

by a board-certified dermatologist or by a physician under their supervision.

Ethics

The study was approved by an independent ethics committee (IBL Life Ltd Ethics Committee, 8 October 2021, study reference no. EC21-COS-067–1). The study complied with the general principles of Good Clinical Practices (GCP) issued by the Helsinki Declaration (and its subsequent modifications) and/or 21 CFR part 50 and the GCP defined by ICH E6(11) ref CPMP/ICH/135/95, 1996, integrated addendum to ICH Topic E6 (R1): Guideline for Good Clinical Practice E6 (R2) current step 4 version dated 9 November 2016, and according to local law.

Statistical Analysis

Efficacy was analysed on an intention-to-treat (ITT) basis, that is, including all subjects who were exposed to the sun. No dropouts were recorded. For subjects with ITA > 41° (exposed for maximum 2 h), the highest values of erythema, pigmentation and corresponding colorimetric readings from the two adjacent sub-areas were used. For those with ITA ≤ 41° (who had one sub-area exposed for 2 h and one for 3 h), the primary analysis used values from the 3-h exposure, that is, the full scheduled exposure. A secondary analysis of all 2-h data was also performed (Supplementary Material).

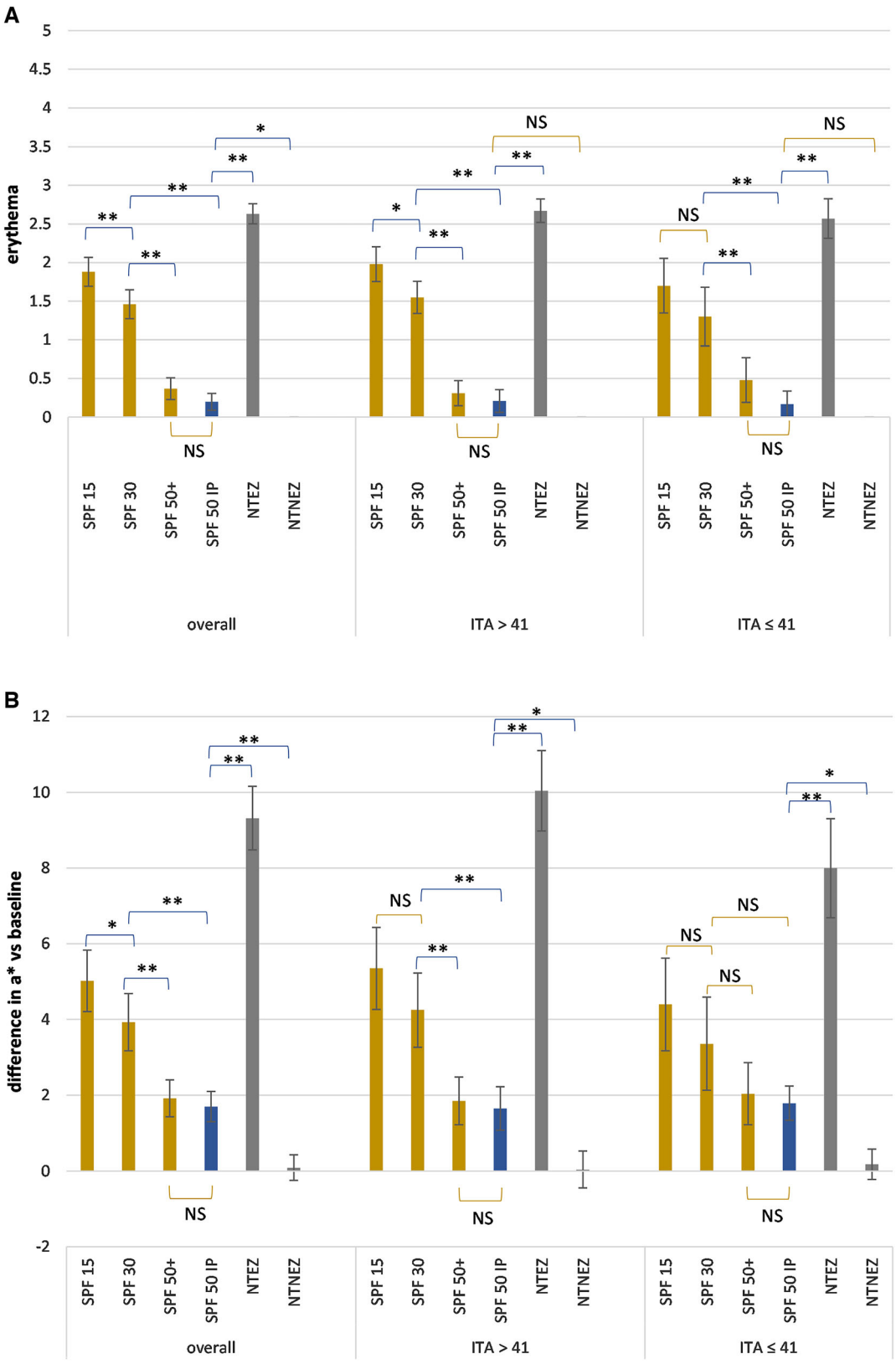
Quantitative variables were reported using measures of central tendency (mean or median) and dispersion (standard deviation). Qualitative variables were reported as count and percentage. Clinical scoring and chromameter parameters were analysed using univariate analysis of variance (ANOVA) with "product" as the fixed factor and "subject" as the random factor, followed by Tukey's procedure for pairwise comparisons. The analysis was conducted on rank-transformed data (equivalent to a non-parametric alternative) for the clinical scores and on the change from baseline for the chromameter measurements. In addition to treating the clinical scores of erythema as quantitative, the frequency of each grade was listed, by product, then grouped into two categories (0–1 and 2–4) to represent photoprotection success and failure, respectively. Statistical significance was set at $p \leq 0.05$. Analysis was performed using SPSS 19.0 and Microsoft Excel 2010 or above.

RESULTS

Figure 2 shows the study population. Sixty-five participants were exposed to the sun and completed the study (ITT population); 40 women and 25 men, mean age 28 years (minimum 18 years, maximum 53 years). Mean ITA was 43.03° [standard deviation (SD) 7.90°, minimum 29.18°, maximum 64.85°]. Forty-two subjects were in the ITA > 41° group, and 23 were in the ITA ≤ 41° group. On the basis of self-reported information, the majority of participants were of white European ethnicity, with a minority (6/65) being of mixed Mauritian-white ethnicity (one parent of Mauritian origin and one parent of European origin); none reported Chinese ethnicity.

UV Radiation Doses

The mean UV index ranged from 5.21 (high) to 10.12 (extreme). The mean cumulative erythemally weighted UVB dose ranged from 83.41 mJ/cm^2 (in four subjects due to a change in weather conditions) to 201.09 mJ/cm^2 after


A
erythema
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
NS
SPF 15
SPF 30
SPF 50+
SPF 50 IP
NTEZ
NTNEZ
overall
ITA > 41
ITA ≤ 41
B
difference in a* vs baseline
12
10
8
6
4
2
0
-2

◀**Fig. 4** Erythema at 24 h. **A**, clinical score; **B**, colorimetry. a* difference versus baseline. The height of the bars indicates the mean, and the error bars indicate the 95% confidence intervals. *$p < 0.05$. **$p < 0.001$. *NS* not significant

2 h and 260.76 mJ/cm^2 after 3 h. The mean cumulative UVA dose ranged from 18.69 J/cm^2 to 40.75 J/cm^2 after 2 h and 54.11 J/cm^2 after 3 h. Six subjects did not receive the minimum cumulative erythemally weighted UVB dose of 100 mJ/cm^2.

During exposure, the mean external temperature was 31.29–35.46 °C with a mean hygrometry between 46.11% and 65.41%.

Clinical Outcomes

A representative photo of one of the participants is shown in Fig. 3, at baseline, 24 h and day 8.

Erythema Clinical Score

At 24 h, the highest erythema scores aligned with the reported SPF values, that is, the products ranked in the appropriate order. Differences in mean clinical erythema score reached statistical significance between adjacent reference product levels: SPF 15 versus 30 and SPF 30 versus SPF 50+. Differences were also significant between the SPF 50 IP and the reference standards SPF 15 and SPF 30 ($p < 0.001$ for each). The difference between the IP and the SPF 50+ reference standard product was not statistically significant ($p = 0.136$). There were also statistically significant differences between the IP and the positive and negative control areas. Figure 4A shows the clinical erythema scores at 24 h, including by skin colour subgroup. These subgroups (light and intermediate skin colour) followed similar patterns to the overall analysis in ranking the products, but comparisons between adjacent protection levels did not always reach statistical significance in these smaller groups.

There was one instance of protection failure (erythema score ≥ 2) with the IP [this individual (RD035) had an ITA > 41° and was exposed for 2 h, to a cumulative erythemally weighted UVB dose of 201.09 mJ/cm^2 and cumulative UVA dose of 37.91 J/cm^2, mean UV index 10.21]; there were four instances of protection failure with the reference standard P8 (SPF 50+) [one individual (RD036) had an ITA > 41° and was exposed for 2 h, to a cumulative erythemally weighted UVB dose of 201.09 mJ/cm^2 and a cumulative UVA dose of 37.91 J/cm^2, and mean UV index of 10.21; the three other subjects had an ITA ≤ 41°, one of whom was exposed for only 2 h (subject RD050), to a cumulative erythemally weighted UVB dose of 192.86 mJ/cm^2 and cumulative UVA dose of 40.75 J/cm^2, with a UV index of 10.50; two (RD022 and RD047) were exposed for 3 h, to, respectively, a cumulative erythemally weighted UVB dose of 202.11 and 192.86 mJ/cm^2 and a cumulative UVA dose of 44.02 and 40.75 J/cm^2 and mean UV index of 7.89 and 10.50]. Table 2 presents the proportion of subjects who fell into the categories of grade 0–1 erythema/successful protection and grade 2–5 erythema/protection failure, by product. No grade 4 or 5 was recorded.

Table 2 Erythema grade at 24 h grouped into photoprotection success and failure

Clinical erythema grade		SPF 15 P3	SPF 30 P5	SPF 50+ P8	SPF 50 IP	NTEZ	NTNEZ
0–1	*N* (%)	19 (29.23)	35 (53.85)	62 (95.38)	64 (98.46)	1 (1.54)	65 (100.00)
2–4	*N* (%)	46 (70.77)	30 (46.15)	3 (4.62)	1 (1.54)	64 (98.46)	0 (0.00)

IP, investigational product; *NTEZ*, non-treated exposed zone, *NTNEZ*, non-treated non-exposed zone; *SPF*, sun protection factor

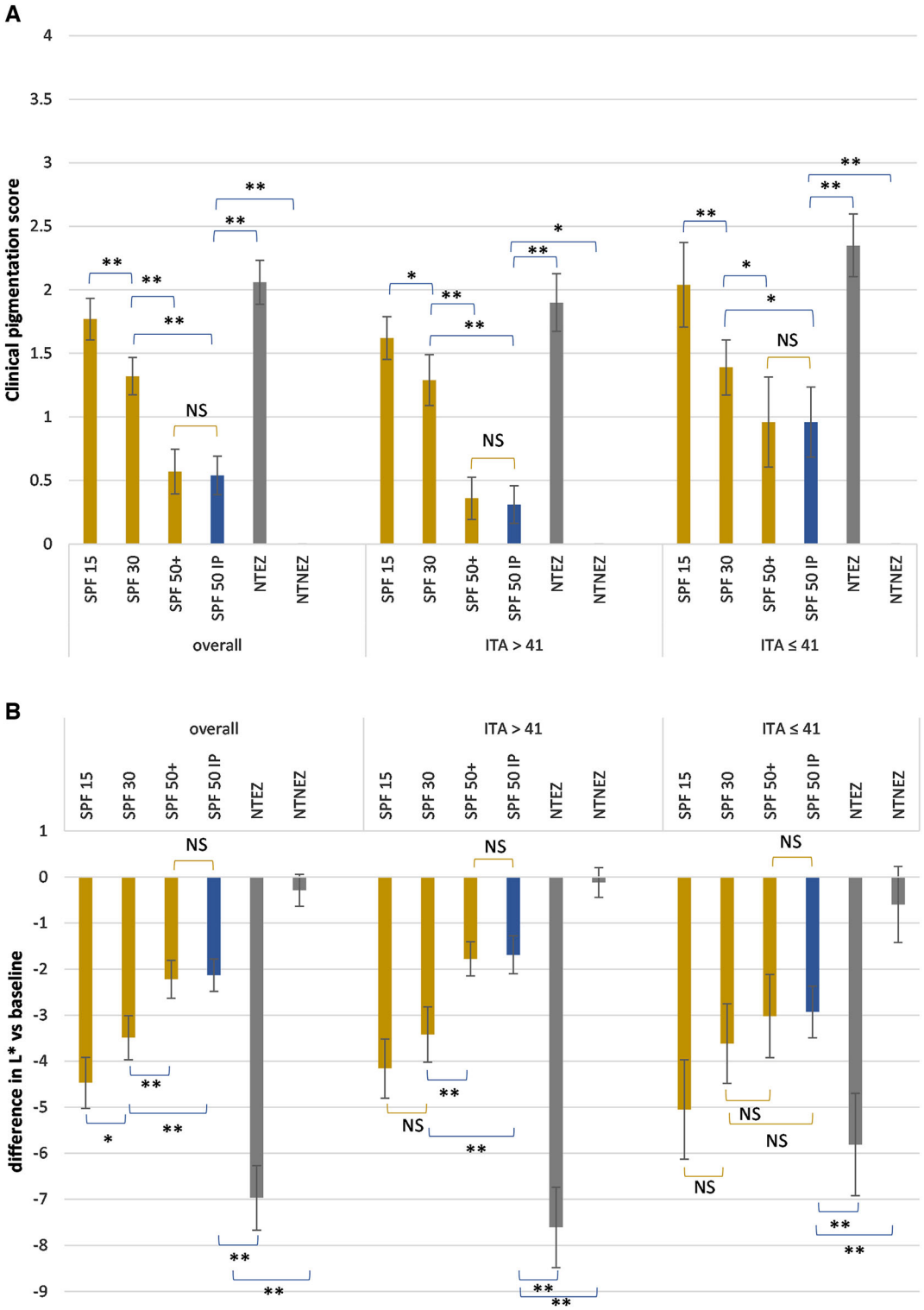


Fig. 5 Pigmentation at day 8. **A** clinical score; **B** colorimetry L*, difference versus baseline; **C** colorimetry ITA, difference versus baseline. The height of the bars indicates the mean, and the error bars indicate the 95% confidence intervals. *$p < 0.05$. **$p < 0.001$. *NS* not significant

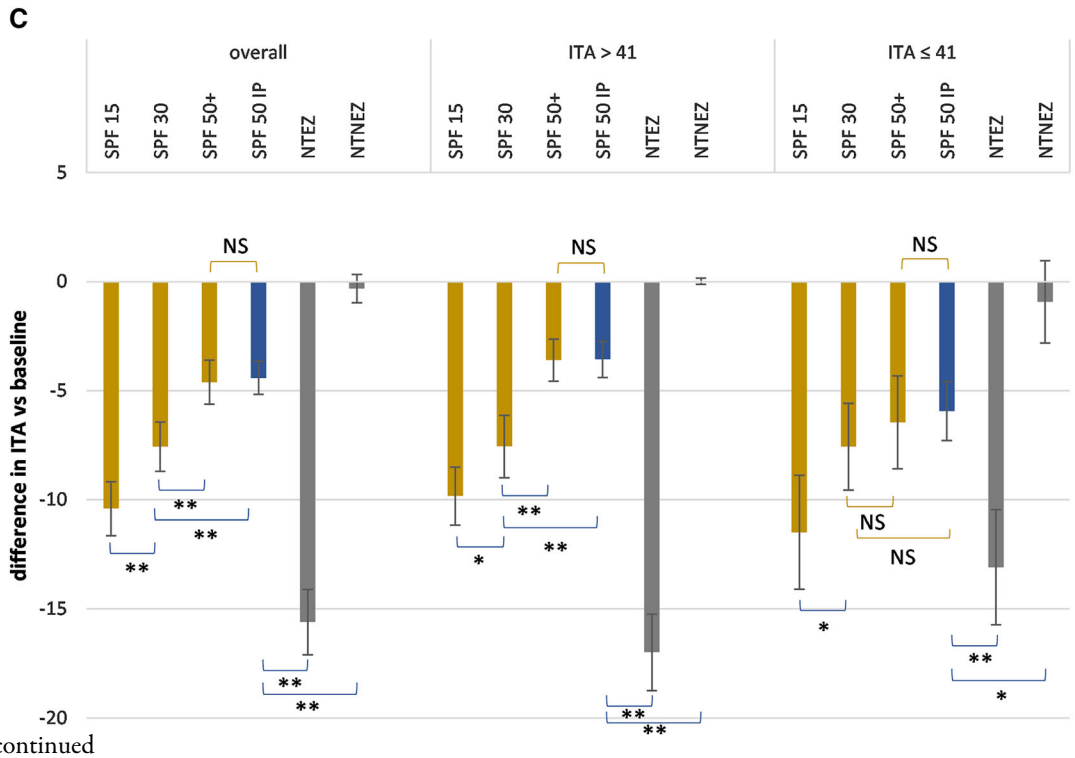


Fig. 5 continued

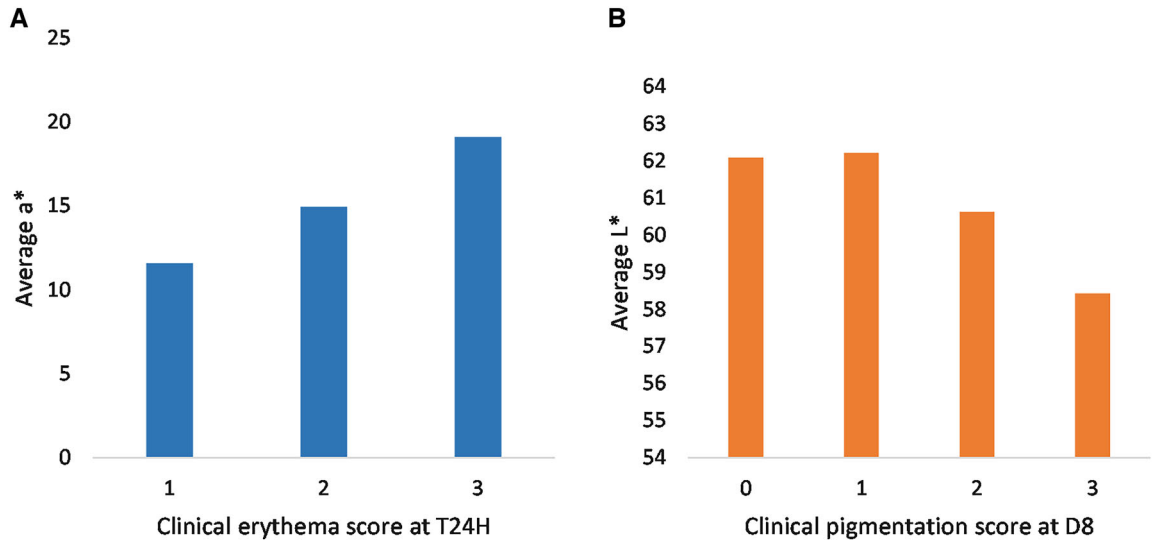


Fig. 6 Correlation between **A** a* and clinical erythema, $P = 0.013$ based on one-way ANOVA, and **B** L* and clinical pigmentation, $P < 0.001$ based on one-way ANOVA

Colorimetry: a*

In line with the clinical erythema scoring, differences in a* from baseline also showed a statistically significant difference between adjacent reference product levels: SPF 15 versus SPF 30 ($p = 0.038$), and SPF 30 versus SPF50+ ($p < 0.001$). The SPF 50 IP was significantly different from the SPF 15 and the SPF 30 ($p < 0.001$

for both), but not from the reference standard SPF 50+ ($p = 0.991$). Again, by skin colour subgroup, similar ranking patterns were observed. Figure 4B shows differences in a* versus baseline at 24 h.

Clinical pigmentation

At day 8, significant differences were present between adjacent protection levels SPF 15 versus 30 and SPF 30 versus 50+ ($p < 0.001$ for both). The SPF 50 IP was significantly different from the SPF 15 ($p < 0.001$) and SPF 30 ($p < 0.001$) but not the SPF 50+ ($p = 1.000$). This was the case overall, and for both light and intermediate subgroups (Fig. 5A).

Colorimetry: L, ITA°*

Overall, L* (difference at day 8 versus baseline) was significantly different for SPF 15 versus SPF 30 ($p = 0.012$) and SPF 30 versus SPF 50+ ($p < 0.001$). The IP was significantly different from SPF 15 and SPF 30 ($p < 0.001$ for both) but not SPF 50+ ($p = 1.000$). The IP was significantly different from both the positive and negative control areas ($p < 0.001$ for both) (Fig. 5B).

For ITA (difference at day 8 versus baseline), the same pattern was seen: statistically significant differences between SPF 15 versus 30 and SPF 30 versus SPF 50+ ($p < 0.001$ for both), and between IP versus SPF 15 and 30 ($p < 0.001$ for both) but not IP versus SPF 50+ ($p = 1.000$). The IP was significantly different from the positive and negative control areas ($p < 0.001$ for both) (Fig. 5C).

The clinical scores correlated well with the colorimetry values, both a* with clinical erythema at 24 h and L* with clinical pigmentation at day 8 (Fig. 6).

Secondary Analyses

Results of the secondary analysis after 2 h of exposure are provided in Supplementary Material. Overall patterns were similar, in that the ranking did not change, but it was not always possible to distinguish statistically significant differences between adjacent SPF levels, especially in the intermediate colour group (though this was a smaller sample).

Reduction of Risk of Severe Burns

In multiple subjects there were areas that had to be covered (27 subjects; 35 areas) before the 2 h period had elapsed owing to immediate grade ≥ 2 erythema. Most of these (23/35) were in the SPF 15 area, and most were in subjects with an ITA > 41° (15/27).

DISCUSSION

The two main findings of this study were, firstly, that it confirmed the capacity of this outdoor clinical efficacy model to discriminate between the three reference standard products (P3/SPF 15, P5/SPF 30 and P8/SPF 50+), supporting the reliability of this method. Secondly, in this model, the new broad-spectrum SPF 50 IP showed better efficacy than reference standards P3 (SPF 15) and P5 (SPF 30) and similar efficacy to reference standard P8 (SPF 50+).

In contrast to previous studies by this group with a similar methodology [11, 12], and therefore somewhat of an unexpected finding, there were also significant differences between the SPF 50 IP area and the unexposed area (negative control), and between the SPF 50+ reference and unexposed area; in our previous studies [11, 12], these high levels of SPF were indistinguishable from the unexposed area. The finding in the present study may be indicative of the strength of the solar exposure in these conditions—the UV index reached a high of 14—and demonstrates the limitations of sunscreens even at high protection levels; covering up was still superior to the highest SPF sunscreens. Indeed, there were instances of sunburn even with these high SPFs, including the SPF 50 IP (one instance for a subject with light skin/ITA > 41°) and the SPF 50+ standard reference product P8 (four instances, one subject with light skin/ITA > 41° and three with intermediate skin colour/ITA ≤ 41°). Clearly this reinforces the importance of not relying solely on sunscreens for protection and avoiding prolonged sun exposure at peak irradiation times; this advice must not be disregarded. It is also curious to note that there were more instances of burn with the SPF 50+ P8 than the SPF 50 IP, although with such small numbers it is difficult

to draw meaningful conclusions on this point. Those cases of burn despite the use of SPF 50+ suggest it may be more appropriate to use an SPF higher than 50 for fair skin when exposed to extreme UV index, particularly when we take into account that users frequently do not apply the recommended 2 mg/cm^2 of sunscreen [13, 14]; in addition, these findings reinforce the recommendation to re-apply sunscreen at least every 2 h and probably more frequently for fair skin. Most importantly, photoprotection with clothes and seeking shade are paramount in such extreme conditions, in addition to avoidance of sun exposure in the peak hours, typically 11 am to 4 pm, whenever possible.

Regarding investigation by other groups in this area, a recent interesting study by Hughes et al. evaluated, in an outdoor setting in Arequipa (Peru), the protection against natural sunlight provided by ten high-SPF (30–110) sunscreens, when exposed to approximately 2 h of natural sunlight [15]. That study, more akin to outdoor SPF testing (despite the fact that natural light is not limited to UVB as in the indoor SPF measurement) than our study, confirmed a significant discrepancy between the actual natural sunlight protection and the SPF and broad-spectrum claims of the sunscreens tested. Like us, Hughes et al. found that SPF correlated well with erythema protection, but the intensity of persistent pigment darkening (PPD) was also higher with higher SPFs. In contrast, in our study, both erythema and pigmentation correlated well (and inversely) with SPF levels. It should be noted that, among other methodological differences, their pigmentation assessments were taken at 24 h, representing PPD, whereas ours were taken at 1 week, representing delayed tanning, so they are not a direct comparison as such. Hughes et al. also noted that 2 h exposure was enough to measure differences in erythema and pigmentation. In the present study, as in a previous study of ours [12], we exposed those with intermediate skin colour for up to 3 h, as we hypothesized that this group may require stronger exposure to allow differentiation between protection levels. In the overall analysis this was true only for one comparison (SPF 15 versus 30, for a*), but in the analysis of the subgroup of intermediate skin colour only, many of the between-level differences were not significant, although the smaller sample size may have been the cause of this. This point will need further monitoring and reflection when planning future studies. We did not find the SPF 50+ P8 to be superior to the SPF 50 IP (in fact, there were more instances of burn with the SPF 50+, though only four in total versus one with the SPF 50 IP, so interpretation of this is limited). A lot has been said about whether to recommend an SPF higher than 30 or 50 [16, 17]: while some argue that the added benefit is low, as the proportion of UV blocked by higher SPFs is relatively small, some investigators have looked at this in outdoor conditions. Williams et al. found, in an outdoor skiing study, that SPF 100 was superior to SPF 50+ for preventing erythema on the face [18], and Russak et al. found a similar pattern of better efficacy beyond SPF 50 [19]. Our findings do not appear to follow this trend, although we used different SPF levels: the SPF 50+ ISO reference product (P8) in our study was determined as having an SPF of 63 in indoor laboratory testing; we cannot rule out that we may have seen a difference had this been an even higher SPF.

Sayre et al. found that protection level in outdoor conditions differed from SPF label, and that this difference was most pronounced when the sun was low in the sky, which was attributed to being due to the changing ratio of UVA to UVB radiation [20]. Our study was carried out close to midday, when UVB is at his highest, but it could be interesting to see if the performance of these products changes with differences in solar height.

Limitations and Strengths

The limitations of this study include the fact that several participants required covering up of test areas before full planned exposure time had elapsed. This may have had an effect on the results, as it would have limited the severity of erythema, but it was done for reasons of safety as allowing them to continue to burn could have been considered a questionable approach. In addition, as the recommendation for

sunscreen use is to re-apply at least every 2 h irrespective of skin colour or phototype, the continuation to 3 h could be criticized. Our rationale for keeping this group exposed for longer than 2 h was that we considered them likely to require stronger exposure to observe significant changes; one alternative would be to include more subjects, although this would also have the potential limitation of introducing more variability between groups, for example with changes in weather conditions on different days of the study. Clearly the conditions of this study differ from those in a laboratory in terms of environmental factors such as temperature control, and, as reported in the results, the mean UV index values ranged from high to extreme; furthermore, the time and costs involved in such a study compared with laboratory methods are not inconsiderable. The reporting of ethnicity was subjective, based on subject-reported descriptions, with the limitations that this entails.

The key strengths are that both clinical and colorimetric assessments appeared to correlate well, in that they ranked the products in the same order, with differences between product levels reaching statistical significance (Fig. 6). It provides additional evidence on the IP and the three reference standard products, confirming that the three reference standard products were correctly ranked in protection level; in these outdoor conditions, we could not differentiate between the SPF 50 IP and the SPF 50+ P8 (with SPF 63). The information provided here contributes to a more holistic view of the products' protection, something which has been called for by leading investigators in the field [21].

Areas for Future Study

Future considerations for improvement of this outdoor clinical efficacy testing model may include modification of the exposure time for the protected areas in order to rank in a statistically significant way the SPF 50 IP and the SPF 50+ reference standard P8.

The model could also be modified regarding the unprotected, exposed area: exposure time for the unprotected area could be reduced to, for example, half an hour, or when the cumulative erythemally weighted UVB dose reaches 50–60 mJ/cm^2, as, at this dose, changes in a* versus baseline were observed for both populations (Supplementary Fig. 3), and all subjects had clinical erythema a score ≥ 1. Indeed, rather than consider duration of sun exposure, we could consider the cumulative dose reached, and this could be different for skin with ITA $>$ 41° and $\leq$ 41°.

The definition of the positive control could be changed to consider grade 1 clinical erythema at 24 h a positive control, rather than grade 2. Some researchers already used this definition in their work, with grade 1 erythema or pigmentation being considered indicative of sun-induced skin injury [15]. However, changing this criterion, and reducing the time of exposure, could carry the risk of some subjects being excluded from the analysis: in our study, only one subject obtained a clinical erythema score of 1 after 1 h. Another option could be to not use an unprotected "positive control" area at all, maintaining the objective of ranking the sunscreens and comparing only the IP versus the reference products SPF 15/P3, SPF 30/P5 and SPF 50+/P8.

CONCLUSION

This study confirms the ability of this outdoor clinical efficacy testing method to rank in order the efficacy of sunscreens tested in conditions of high to extreme natural sun exposure from P3 (SPF 15) to P8 (SPF 50+) and to position a new broad-spectrum SPF 50 IP as no different to the reference standard P8 (SPF 50+) and superior to P5 (SPF 30). This information on performance in natural sun exposure is intended to be taken as supplementary to the established laboratory-determined values. The instances of burn even with high levels of protection underscore the need for topical photoprotection to be combined with other established protective behaviours such as seeking shade and using protective clothing, hats and sunglasses.

ACKNOWLEDGEMENTS

We thank all the volunteers who participated in this study. Thanks also to Alba Villanueva, clinical manager at ISDIN, for help in managing this study.

Funding. This study and the journal's Rapid Service Fee were funded by ISDIN.

Medical Writing Assistance. Drafting and editorial assistance in the preparation of this article was provided by Jane Marshall of Marshall Medical Translation. Support for this assistance was funded by ISDIN.

Author Contributions. CG designed the protocol, validated the data analysis and the final report of the clinical study, drafted the outline of the publication and reviewed and validated the content of the publication. GP, as clinical investigator, selected the subjects, supervised them during the sun exposure to reduce the risk of sunburn, performed the clinical evaluation, and reviewed and validated the content of the publication. MH designed the statistical analysis, supervised the quality and the coherence of the data and the final report, and reviewed and validated the content of the publication. JB, as clinical manager, validated the protocol, managed the operational aspects of the clinical study, and validated the clinical report and the content of the publication. CT validated the protocol, the statistical analysis and the final clinical study report and the content of the publication. TP performed critical review of the manuscript for intellectual content. JK performed critical review of the manuscript for intellectual content.

Disclosures. Corinne Granger was employed by ISDIN when the protocol was developed and the clinical study carried out. She is currently appointed as a consultant by ISDIN. Gitanjali Petkar is employed by CIDP, the clinical research organization that ran the clinical study. Muzzammil Hosenally is a consultant appointed by CIDP, the clinical research organization that ran the clinical study. Javier Bustos is an employee of ISDIN. Carles Trullàs is an employee of ISDIN. Thierry Passeron is a consultant appointed by ISDIN. Jean Krutmann is a consultant appointed by ISDIN.

Compliance with Ethics Guidelines. The study was approved by an independent ethics committee (IBL Life Ltd Ethics Committee, 8th October 2021, study reference no. EC21-COS-067–1). The study complied with the general principles of Good Clinical Practices issued by the Helsinki Declaration (and its subsequent modifications) and/or 21 CFR part 50 and the GCP (Good Clinical Practices) defined by ICH E6(11) ref CPMP/ICH/135/95, 1996, Integrated addendum to ICH Topic E6 (R1): Guideline for Good clinical Practice E6 (R2) current step 4 version dated 9 Nov 2016 and new cosmetic regulation. All participants gave signed informed consent to take part including use of photographs in publications.

Data Availability. The datasets generated during and/or analyzed during the current study are not publicly available due to confidential information on the IP.

REFERENCES

1. Geisler AN, Austin E, Nguyen J, Hamzavi I, Jagdeo J, Lim HW. Visible light. Part II: photoprotection against visible and ultraviolet light. J Am Acad Dermatol. 2021;84(5):1233–44

2. Organization IS. Cosmetics—sun protection test methods. In vivo determination of the sun protection factor (SPF). 2019.

3. Organization IS. Cosmetics—sun protection test methods. In vivo determination of sunscreen UVA protection. 2011.

4. Organization IS. ISO 24443:2021. Cosmetics—determination of sunscreen UVA photoprotection in vitro. 2021.

5. Surber C, Osterwalder U. Challenges in sun protection. Curr Probl Dermatol. 2021;55:1–43.

6. Diffey B, Osterwalder U. Labelled sunscreen SPFs may overestimate protection in natural sunlight. Photochem Photobiol Sci. 2017;16(10):1519–23.

7. Kohli I, Chaowattanapanit S, Mohammad TF, Nicholson CL, Fatima S, Jacobsen G, et al. Synergistic effects of long-wavelength ultraviolet A1 and visible light on pigmentation and erythema. Br J Dermatol. 2018;178(5):1173–80.

8. Kohli I, Zubair R, Lyons AB, Nahhas AF, Braunberger TL, Mokhtari M, et al. Impact of long-wavelength ultraviolet al and visible light on light-skinned individuals. Photochem Photobiol. 2019;95(6):1285–7.

9. Miksa S, Lutz D, Guy C, Delamour E. Sunscreen sun protection factor claim based on in vivo interlaboratory variability. Int J Cosmet Sci. 2016;38(6):541–9.

10. Osterwalder U, Surber C. Characterization of sun protection performance: quo vadis? Hautarzt. 2022;73(4):276–82.

11. Granger C, Krutmann J, Bustos J, Sola Y, Hosenally M, Trullàs C, et al. New methodology to evaluate sunscreens under outdoor conditions: a double-blind, randomized intra-individual clinical study of a water-based broad-spectrum SPF50+ versus SPF15 (P3) and SPF50. Dermatol Ther (Heidelb). 2019;9(3):589–99.

12. Granger C, Ong G, Andres P, Trullas C, Hosenally M, Lai W, et al. Outdoor sunscreen testing with high-intensity solar exposure in a Chinese and Caucasian population. Photodermatol Photoimmunol Photomed. 2021.

13. Azurdia RM, Pagliaro JA, Diffey BL, Rhodes LE. Sunscreen application by photosensitive patients is inadequate for protection. Br J Dermatol. 1999;140(2):255–8.

14. Bech-Thomsen N, Wulf HC. Sunbathers' application of sunscreen is probably inadequate to obtain the sun protection factor assigned to the preparation. Photodermatol Photoimmunol Photomed. 1992;9(6):242–4.

15. Hughes SNG, Lowe NJ, Gross K, Mark L, Goffe B, Hughes H, et al. Assessment of natural sunlight protection provided by 10 high-SPF broad-spectrum sunscreens and sun-protective fabrics. Curr Probl Dermatol. 2021;55:157–69.

16. Farberg AS, Glazer AM, Rigel AC, White R, Rigel DS. Dermatologists' perceptions, recommendations, and use of sunscreen. JAMA Dermatol. 2017;153(1):99–101.

17. Herzog SM, Lim HW, Williams MS, de Maddalena ID, Osterwalder U, Surber C. Sun protection factor communication of sunscreen effectiveness: a web-based study of perception of effectiveness by dermatologists. JAMA Dermatol. 2017;153(3):348–50.

18. Williams JD, Maitra P, Atillasoy E, Wu MM, Farberg AS, Rigel DS. SPF 100+ sunscreen is more protective against sunburn than SPF 50+ in actual use: results of a randomized, double-blind, split-face, natural sunlight exposure clinical trial. J Am Acad Dermatol. 2018;78(5):902-10.e2.

19. Russak JE, Chen T, Appa Y, Rigel DS. A comparison of sunburn protection of high–sun protection factor (SPF) sunscreens: SPF 85 sunscreen is significantly more protective than SPF 50. J Am Acad Dermatol. 2010;62(2):348–9.

20. Sayre RM, Kollias N, Ley RD, Baqer AH. Changing the risk spectrum of injury and the performance of sunscreen products throughout the day. Photodermatol Photoimmunol Photomed. 1994;10(4):148–53.

21. Surber C, Uhlig S, Bertrand C, Vollhardt J, Osterwalder U. Past, present, and future of sun protection metrics. Curr Probl Dermatol. 2021;55:170–87.