J. Cosmet. Sci., **70,** 1–15 (January/February 2019)

Evaluating the Moisturizing Abilities and Sun Protection Factor of New Lip Balm Formulations

CHRISTOPH GFELLER, GEORGE HARDIE,
GILBERT SHANGA, and HARISH MAHALINGAM,
R&D Medical Affairs Skin Health, GSK Consumer Healthcare, Brentford, Middlesex TW8 9GS, United Kingdom (C.G.), R&D Clinical Operations, GSK Consumer Healthcare, Weybridge, Surrey KT13 0DE, United Kingdom (G.H.), R&D Biostatistics, GSK Consumer Healthcare, Warren, New Jersey (G.S.), R&D Medical Affairs Skin Health, GSK Consumer Healthcare, Warren, New Jersey (H.M.)

Accepted for publication September 12, 2018.

Synopsis

This report explores dry-skin models to assess the potential of a new lip balm formulation to hydrate dry skin or lips, and presents sun protection factor (SPF) values for five new lip balm formulations. Evaporimeter [for transepidermal water loss (TEWL)], Skicon®, and Corneometer® were used to measure hydrating effects of lip balm formulations in a dry-skin leg model, and TEWL, DermaLab® Moisture Meter, Corneometer®, and visual assessments were used with a dry-lip model. SPF studies were conducted in accordance with either the U.S. Food and Drug Administration monograph final rule or international standard ISO 24444. Data from dry-skin leg model demonstrate that a new lip balm formulation significantly improves skin hydration compared with untreated leg skin and four comparator products. Data obtained from a dry-lip model proved unreliable. Five new lip balm formulations exhibited sunscreen capability; however, they did not meet the intended SPF. There were no product-related adverse events with the formulations. Although the new lip balm formulation improved hydration, data from a novel dry-lip model proved unreliable therefore further testing is required to confirm these benefits. Five new lip balm formulations provided sunscreen capability but did not meet the intended SPF, and will undergo reformulation and retesting.

INTRODUCTION

Exposure to small amounts of ultraviolet (UV) radiation is known to be beneficial, although prolonged UV exposure may be damaging for the skin, eyes, and immune system (1). Erythema, commonly referred to as sunburn, is an acute damaging effect of UV exposure on the skin (2). Prolonged UV exposure, however, may cause degenerative changes in the skin and fibrous tissues, such as drying or coarsening of the skin, loss of elasticity, and premature skin aging (2). Moreover, prolonged exposure to UV radiation can also induce

Address all correspondence to Christoph Gfeller at christoph.f.gfeller@gsk.com.

permanent skin pigmentation, such as freckles and moles, and actinic keratosis, and may significantly increase the risk of nonmelanoma skin cancer (NMSC) (2,3).

Restricting exposure to the sun combined with regular use of sunscreen products with an appropriate sun protection factor (SPF) can limit the damaging effects of UV rays, including the risk of NMSC (1). SPF is an important measure of efficacy for sunscreen products, providing an international standard for the level of protection provided against erythema induced by ultraviolet B (UVB) radiation (1).

Sunscreen products containing active photoprotectant ingredients must undergo rigorous regulatory testing to ensure safety and effectiveness before approval is granted (4). This includes *in vivo* UVB testing in accordance with either the U.S. Food and Drug Administration (FDA) monograph final rule requirements (5) or international standard ISO 24444 (6).

In addition to safety and effectiveness in the context of SPF, another important consideration for the development of new lip balm formulations is their moisturizing abilities, *i.e.*, the reduction of water loss from the underlying tissue. The stratum corneum (SC), a lipid matrix in the outermost layer of the skin, is crucial for skin permeability barrier function and acts to prevent excessive water loss through the skin. One of the most important parameters for determining SC barrier function is transepidermal water loss (TEWL), the loss of water by passive evaporation through the skin (7–9). Separately, skin hydration can be measured using a variety of techniques, and instruments such as Skicon® (I.B.S. Co., Hamamatsu, Japan), DermaLab® (Cortex Technology, Hadsund, Denmark) Moisture Meter (DMM), and Corneometer® (Courage + Khazaka Electronic GmBH, Cologne, Germany) are considered the industry standards for this purpose (10).

Newer lip balm formulations have been developed to not only provide protection from the sun, but also with the potential to hydrate dry lips. These new lip balm formulations include emollients and lipids combined with glycerin, and photoprotectant compounds that provide defined SPF. To meet regulatory requirements and obtain approval for use, rigorous clinical testing of these new lip balm formulations is required. Data from such clinical studies may also support product claims, including those pertaining to potential skin hydration and/or moisturizing abilities.

Here we present data from five clinical studies evaluating a series of novel lip balm formulations. The aim of four of these studies was to determine the SPF of five new lip balm formulations to comply with sun protection labeling as defined by the U.S. FDA monograph final rule (5) and international standard ISO 24444 (6). The aim of the fifth study was to develop a new *in vivo* methodology to help differentiate lip balm formulations that have the potential to hydrate dry lips. In this study, the moisturizing abilities of one of the new lip balm formulations was first compared with four comparator/reference lip products using a dry-skin leg model and evaporimeter (TEWL), Skicon, and Corneometer methodologies. Subsequently, an exploratory, home-use phase was conducted to assess the moisturizing abilities of this lip balm formulation over time using a novel dry-lip model and similar skin hydration assessment methodologies.

MATERIALS AND METHODS

STUDY DESIGN AND CONDUCT

Barrier function and moisturizing abilities study. Study RH02116 was a randomized, six-cell, block design study to evaluate the effect on barrier function and moisturizing abilities of

a new lip balm formulation, test product B, in comparison with four reference lip balm formulations: comparator A, Aquaphor® Lip Repair + Protect Broad Spectrum SPF 30 (Beiersdorf, Hamburg, Germany); comparator B, Blistex Five Star Lip Protection® SPF 30 (Blistex, Oak Brook, IL); comparator C, Carmex® Everyday Protecting Lip Balm Stick SPF 15 (Carmex, Franklin, WI); and comparator D, Neosporin® Lip Health™ Daily Hydration Therapy Sunscreen SPF 20 (Johnson & Johnson Consumer, Inc., New Brunswick, NJ).

The study was conducted between November 2013 and January 2014 at cyberDERM Clinical Studies, Broomall, PA. The study was approved by the institutional review board and was conducted in accordance with Good Clinical Practice (GCP) guidelines and the ethical principles of the Declaration of Helsinki. The investigator and site staff were responsible for identifying, documenting, and reporting adverse events (AEs).

SPF determination studies. Four studies were conducted to determine the SPF value of a total of five new lip balm formulations. All SPF studies were evaluator-blind, randomized, complete-block design trials. The SPF standard P2 (7% padimate O, 3% oxybenzone) positive control, with an accepted SPF of 16, was included in the four SPF studies to confirm the accuracy of the protocols being used.

Study RH01927 was conducted between July 2013 and August 2013, whereas study RH01928 was conducted in August 2013. Study RH02117 was conducted between October 2013 and November 2013. These three studies were conducted at the TKL Research, Inc., Paramus, NJ. The fourth study, RH02385, was conducted between March 2014 and April 2014 at proDERM Institute, Hamburg, Germany.

Studies RH01927 and RH01928 determined the SPF of a lip balm formulation, test product A, in accordance with the FDA monograph final rule (5) and ISO 24444 (6), respectively. The protocols were approved by the IntegReview Ethical Review Board.

Study RH02117 determined the SPF of two lip balm formulations, test product B and test product C, in accordance with the FDA monograph final rule (5), with the approval of the study protocol by the IntegReview Ethical Review Board.

Study RH02385 determined the SPF of two further lip balm formulations, test product D and test product E, according to both the FDA monograph final rule (5) and ISO 24444 (6). The protocol received approval from the Freiburg Ethics Commission International.

All studies were conducted in accordance with GCP guidelines and all participants provided written informed consent in accordance with the requirements of the Declaration of Helsinki. The investigators and site staff for each of these studies were responsible for identifying, documenting, and reporting any AEs.

PARTICIPANTS

Barrier function and moisturizing abilities study. Eligibility criteria for participants in study RH02116 included the following: females aged 18–45 years who are self-identified dry-leg sufferers and scored between 2 and 4 on the nine-point dry-leg scale (0–4 in increments of 0.5) on visual inspection (11,12). Exclusion criteria included marks, scars, scratches, tattoos, or other blemishes on the test site; a history of active eczema, psoriasis, ichthyosis, or any other skin condition; an intolerance or hypersensitivity to the study materials; and use of antihistamines 3 days before visit 1 or during the course of the study.

SPF determination studies. For studies RH01927, RH01928, and RH02117, eligibility criteria included the following: aged ≥18 years; uniformly colored skin on the lower thoracic area of the back; Fitzpatrick skin type I, II, or III (13); and free of any systemic or dermatological disorder. Exclusion criteria included the following: any visible skin disease, excessive hair, blemishes or moles, tan, dermal lesions, or uneven pigmentation on the back; a medical history of skin cancer; a known sensitivity to cosmetics, skin care products or topical drugs related to the study material; and using a medication suspected of causing photobiological reactions, or an abnormal sunlight response.

For study RH02385, eligibility criteria included the following: aged 18–70 years; uniformly colored skin; no erythema or dark pigmentation on the test site; Fitzpatrick skin type I, II, or III (13,14); and an individual typology angle (ITA) > 28° on the test site (15). Exclusion criteria included moles, tattoos, scars, irritated skin, hair, or active skin disease on the test site; a medical history of dysplastic nevi or melanoma; a known sensitivity to cosmetics, skin care products or topical drugs related to the study material; using a medication suspected of causing photobiological reactions within 14 days before the study starts; and a history of abnormal response to sunlight.

RANDOMIZATION AND MASKING

Barrier function and moisturizing abilities study. Study RH02116 consisted of two phases: a leg hydration phase comprising a single application of test product on a single day; and an exploratory home-use, lip-hydration phase comprising six applications of test product daily over an 11-d period.

Phase 1: Dry-skin leg model. In the leg hydration phase, three test sites were marked on each calf of each participant, for a total of six test sites. Test product B was applied to the first test site, the four comparator lip balm products were each applied to a different site, and the sixth site was an untreated control. The study investigators were blinded to treatment allocation at each test site.

During the leg hydration phase, participants visited the clinic twice over a 7-d period. At visit 1, participants were screened and visually assessed for leg dryness using a nine-point dryness scale [0–4, measured in increments of 0.5, where 0 = none, 1 = slight flaking/uplifting of flakes (patchy and/or powdered appearance), 2 = moderate flaking/uplifting flakes [uniform] and/or slight scaling, 3 = severe flaking/scaling, uplifting of scales and/or slight fissuring, 4 = severe scaling/uplifting scales; with severe fissuring/cracking] (11,12). This was followed by a dry-down period of 5–7 days during which participants could not use moisturizing products on their legs. At visit 2, after the dry-down period, participants were reassessed for leg dryness and enrolled if they exhibited grade ≥2 on the leg dryness scale. Assessments were carried out before product application, and at 1, 3, and 6 hours after product application.

Phase 2: Home-use lip-hydration phase. In the home-use, lip-hydration phase of the study, a subset of participants who completed the leg hydration phase were screened for lip dryness using a nine-point scale (0–8, where 0 = no dryness or chapping evident, 1–2 = slight, but definite roughness; fine scaling, 3–4 = moderate roughness; coarse scaling; slight cracking, 5–6 = marked roughness; coarse scaling; obvious cracking, 7–8 = very marked roughness; coarse scaling; cracked progressing to fissuring) (16). Qualifying

participants exhibited lip dryness of grade 3 or 4 (16). The participants then underwent a dry-down period of 3–7 days during which the use of moisturizing products on their lips was prohibited. This was followed by a baseline visit and five assessment visits over an 11-d period (days 3, 5, 8, 10, and 12). The participants applied test product B six times daily between the five assessment visits.

SPF determination studies. In studies RH01927 and RH01928, four test sites were selected on each participant's back and randomized to one of four treatments: two sites were untreated negative controls, the SPF standard P2 positive control was applied to a third test site, and test product A was applied to the fourth test site. Study investigators were blinded to the treatment allocation for each of the test sites.

The four test sites were each divided into six subsites for UV irradiation. Participants attended the study clinic on three consecutive days. At visit 1, all six subsites of one untreated control test site were exposed to UV light. Irradiated sites were visually assessed at visit 2 using the erythema grading criteria (Table I) (17) to detect the minimal erythemal dose (MED) of the unprotected site (MEDu).

Treatment application (where appropriate) and UV irradiation at the remaining untreated control test site and two treatment test sites were conducted at visit 2. Test product A and the SPF standard P2 positive control were applied using a finger cot and allowed to dry for ≥15 min before UV irradiation. At visit 3, the effect of UV irradiation on the various test sites was assessed using the erythema grading criteria and the MED determined for test product A.

In study RH02117, five test sites were selected on each participant's back and randomized to one of five treatments: two sites were untreated negative controls, the SPF standard P2 positive control was applied to a third test site, and test products B and C were applied to a fourth and fifth test site, respectively. Study investigators were blinded to the treatment allocation for each of the test sites. Each test site was divided into five subsites before UV irradiation.

The participants attended the study clinic on three consecutive days. At visit 1, all five subsites of one untreated control test site were exposed to UV light, using a Xenon Arc Solar Simulator2 (150 W) (Solar Light Co., Inc., Philadelphia, PA). Irradiated sites were visually assessed at visit 2 using the erythema grading criteria (Table I) (17) to determine MEDu. Treatment application and UV irradiation at the remaining four test sites were conducted at visit 2. Test products B and C and the SPF standard P2 positive control were applied using a finger cot and allowed to dry for 30 min before UV irradiation. At visit 3, the effect of UV irradiation on the various test sites was assessed using the erythema grading criteria, and the MED was determined for test product B and test product C.

Table I
Erythema Grading and Definitions (17)

Grade	Definition
0	No reaction
1	Minimal or doubtful erythema, barely perceptible compared with the surrounding
2	Mild, but definite erythema with clearly defined borders
3	Moderate erythema
4	Marked/severe erythema

In study RH02385, four test sites were selected on each participant's back and randomized to one of four treatments: one site was an untreated negative control, the SPF standard P2 positive control was applied to a second test site, and test products D and E were applied to the third and fourth test sites, respectively. The study investigators were blinded to the treatment allocation for each of the test sites. Each test site was divided into six subsites for UV irradiation.

The participants attended the study clinic on 4 d. At visit 1, all six subsites of the untreated control site were exposed to UV light, using a 300 W Xenon arc lamp (Solar Light Co., Inc., Philadelphia, PA). Irradiated sites were visually assessed at visit 2, the day after visit 1, using the erythema grading criteria (Table I) (17) to determine MEDu. Treatment application and UV irradiation at the remaining three test sites were conducted at visit 3, within a week of visit 1. Test products D and E and the SPF standard P2 positive control were applied using a finger cot and allowed to dry for 15–30 min before UV irradiation. At visit 4, within 16–24 h of visit 3, the effect of the UV irradiation on the various test sites was assessed using the erythema grading criteria and the MED determined for test product D and test product E.

STUDY ASSESSMENTS

Barrier function and moisturizing abilities study. In phase 1 of study RH02116, the effect of a new lip balm formulation and four comparator lip balm products on barrier function was determined using TEWL, assessed using a cyberDERM RG1 evaporimeter, as a measure of barrier function in a dry-skin leg model. The moisturizing abilities of the five lip balm formulations were also assessed in the dry-skin leg model, using visual assessment, Skicon, and Corneometer methods. Baseline barrier function and skin hydration were assessed before a single application of each lip balm, and at 1, 3, and 6 h postapplication.

In phase 2 of study RH02116, TEWL, DMM, Corneometer, and visual assessment were used to measure changes in the moisture content of the lips after application of test product B. Visual assessment of lip dryness was carried out by a trained evaluator: 0 was categorized as no dryness and 4 was categorized as severe scaling with severe fissuring or cracking. Test product B was applied six times daily by the participants over an 11-d period.

SPF determination studies. For the SPF studies, the MED for each participant was determined for an unprotected test site (MEDu) as a negative control, a site treated with the test product (tpMEDp), and a test site treated with the SPF standard control product (ssMEDp).

MED is defined as the smallest UV dose that produces perceptible redness of the skin with clearly defined borders 20–24 h after UV irradiation using erythema grading criteria (17) (Table I). A series of six UV doses increasing by increments of 25% were administered to the six test subsites of the untreated site. A series of six UV doses increasing by increments of 20% were administered to the six test subsites of the sites treated with test products or the P2 control. In all UV dose series, the middle dose administered was equivalent to MEDu multiplied by the expected SPF of the test product on untreated skin.

Exposure time was influenced by the participant's skin phototype and the power of the solar simulator. MED was measured through visual observation of the test site immediately after irradiation and 16–24 h postirradiation. In study RH02385, the mean ITA of six repetitive chroma meter measurements was calculated to categorize the skin color and skin type of each participant (15).

STUDY OUTCOMES

Barrier function and moisturizing abilities study. The primary endpoint for study RH02116 was to determine the effect of five lip balms (one test formulation and four comparator products) on barrier function 1, 3, and 6 h postapplication using the dry-skin leg model. The secondary endpoint was to determine the moisturizing abilities of the five lip balms 1, 3, and 6 h postapplication, using the dry-skin leg model. Additional exploratory objectives included the following: to explore the effect of test product B on barrier function on dry lips during home use; to explore the moisturizing abilities of test product B on dry lips using DMM, Corneometer, and clinical visual assessment of lip dryness during home use.

SPF determination studies. The primary efficacy endpoint for the four SPF studies was the individual SPF value (SPFi) of the test lip balm formulations 20–24 h after UV irradiation, determined by SPFi = MEDp/MEDu.

An additional primary endpoint for studies RH02117 and RH01927 was label SPF, defined as the largest whole number less than SPF – (t* SE), where t is taken from the subjects' t distribution table corresponding to the upper five percent point with n – 1 degrees of freedom and SE is the standard error of the SPF value.

The secondary endpoints were the safety and tolerability of the lip balm formulations.

SAFETY ASSESSMENT

In all five studies, AEs were defined as any untoward medical occurrence in a patient or clinical investigation subject temporally associated with the use of an investigational product, whether or not considered related to the investigational product. Serious AEs (SAEs) were defined as any AE that results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect. In study RH02116, AEs and SAEs were collected from the start of investigational product use until 5 d following last application of the investigational product. In the SPF studies, AEs and SAEs were recorded from the start of UV irradiation until 5 days following last application of the investigational product. SAEs related to study participation were recorded from the time the participant consented up to and including any follow-up contact. Subjects were asked about AEs at each study visit. In study RH02116, AEs were listed by treatment phase. Prior and concomitant medication use and concomitant nondrug treatment/procedures were recorded for each participant in all five studies. Assessment of the relationship between the investigational product and the occurrence of each AE or SAE was performed by the respective investigator.

STATISTICAL ANALYSES

Barrier function and moisturizing abilities study. Study RH02116 was an exploratory study and no formal sample size calculation was undertaken; however, 20 participants was considered adequate. All valid data from each of the four objective measurements (TEWL, Skicon, DMM, and Corneometer) of either barrier function or skin hydration were used in the intent-to-treat (ITT) analysis. All participants enrolled in this study were also considered evaluable for the determination of SPF. Statistical analyses included calculation of the mean,

standard deviation (SD), median, minimum, and maximum for each assessment parameter for each of the products evaluated, in both the leg hydration phase and the home-use lip hydration phase of the study. Analysis of covariance (ANCOVA) was used to compare test product B with each of the four comparator products at each time point using TEWL, Skicon, and Corneometer measurements. The ANCOVA model included treatment as a fixed effect, subject as a random effect, and baseline values on a site level and subject level for each measurement as covariates.

SPF determination studies. In the SPF studies, data were rejected if irradiation did not achieve an MED or an erythemal response, if all subsites showed an erythemal response, or if responses on the treatment site were absent 20–24 h postirradiation. The calculated SPF for a test product was considered valid if the SPF value of the positive control P2 fell within the SD range of the intended SPF, for either ISO 24444 (16.4 ± 2.4) or the FDA final rule (16.3 ± 3.43). No formal statistical comparisons of calculated SPF values between the five test products were conducted.

In studies RH01927 and RH02117, sample size calculations determined that a minimum of 10 participants, and no more than 13, were required per product to produce valid test results for a screening study. Statistical analyses included calculation of the mean, SD, coefficient of variation (CV), and lower one-sided 95% confidence interval (CI) of the SPF for each product.

In study RH01928, sample size calculations determined that a minimum of 10 participants and a maximum of 25 were required per product. A minimum of 10 valid participant SPF results and a maximum of 20 were required to determine the mean SPF of each product. The number of valid SPF results was deemed sufficient when the 95% CI was within ±17% of the mean SPF. If more than 20 results were needed to reach this criterion, the test was deemed invalid. Statistical analysis included calculation of the mean, SD, CV, SE, and 95% CI of the SPF for each product.

In study RH02385, the safety population (SP) used for the safety analysis included all enrolled participants who received the first irradiation for MEDu determination. The per-protocol population used for efficacy analysis was defined per product and included all participants meeting the inclusion and exclusion criteria and who completed the study without protocol violations. A minimum of 10 participants were recruited as per the FDA final rule (5) and ISO 24444 (6), with additional participants recruited if required for ISO 24444, up to a maximum of 25. The statistical analyses included calculation of the mean and 95% CI of the SPF for each product.

RESULTS

DISPOSITION AND BASELINE CHARACTERISTICS

Barrier function and moisturizing abilities study. In study RH02116, a total of 39 individuals were screened, of whom 24 were randomized and included in the SP and ITT populations; 12 participants did not meet the study criteria and three participants decided not to participate before randomization. One individual did not complete the study because of failure to meet baseline criteria. All participants were female Caucasians with a mean age of 38.2 years (Table II). A subset of participants ($n = 4$) were enrolled into the home-use lip hydration phase of the study.

SPF determination studies. Ten participants were enrolled into study RH01927, all of whom completed the study. All 10 were Caucasian, the majority (8/10) were female, and the mean age was 53.6 years. All participants were classified with a skin type of Fitzpatrick grade II (6/10) or III (4/10) (Table II).

Sixteen participants were enrolled into study RH01928 with no study discontinuations. All were Caucasian and the majority (68.8%) were female. The mean age was 57.0 years. All participants were classified with a skin type of Fitzpatrick grade II (68.8%) or III (31.3%) (Table II). Data from three participants were deemed invalid because of a lack of protection at all UV-irradiated subsites following application of test product A. These participants were therefore excluded from the analysis.

In study RH02117, 16 participants were enrolled. Three participants discontinued before product testing because of study ineligibility ($n = 2$) or because the study had recruited sufficient numbers ($n = 1$). All participants were Caucasian, the majority (87.5%) were female, and the mean age was 55.1 years. Skin type for all participants was either Fitzpatrick grade II (37.5%) or III (62.5%) (Table II). For test product B, data from four of the participants were deemed invalid either because of all subsites showing an erythemal response 20–24 h after irradiation ($n = 3$) or because of a lack of erythemal response 20–24 h after UV irradiation ($n = 1$). Similarly, for test product C, data from five of the participants were deemed invalid, either because of all ($n = 4$) or no ($n = 1$) subsites showing an erythemal response 20–24 h after irradiation.

Thirteen participants were screened and enrolled into study RH02385 to meet the FDA monograph final rule requirements. An additional three participants were enrolled to meet ISO 24444 study requirements, giving a total sample size of 16. Of the 13 who were enrolled for the FDA protocol testing of test products D and E, data from three participants were excluded from the analysis of each test product, either due to a lack of erythema MED ($n = 2$) or invalid data for either the P2 positive control or the test product ($n = 1$). For the same reasons, data for these three participants were also excluded from the 16 who were enrolled for the ISO 24444 protocol testing of test product D.

Of the 16 participants who were enrolled for the ISO 24444 protocol testing of test product E, data from four were excluded because of the lack of erythema MED ($n = 2$); invalid

Table II
Baseline Characteristics of Participants in the Five Studies

	RH02116, $n = 24$	RH01927, $n = 10$	RH01928, $n = 16$	RH02117, $n = 16$	RH02385 (FDA), $n = 13$	RH02385 (ISO 24444), $n = 16$
Mean age, range (years)	38.2 (24–45)	53.6 (44–66)[a]	57.0 (43–70)[a]	55.1 (35–70)[a]	41.2 (20–59)	42.2 (20–59)
Female, *n* (%)	24 (100)	8 (80.0)	11 (68.8)	14 (87.5)	9 (69.2)	12 (75.0)
Fitzpatrick grade, *n* (%)						
I	NA	0 (0)	0 (0)	0 (0)	4 (30.8)	6 (37.5)
II	NA	6 (60.0)	11 (68.8)	6 (37.5)	7 (53.8)	8 (50.0)
III	NA	4 (40.0)	5 (31.3)	10 (62.5)	2 (15.4)	2 (12.5)
Mean ITA,	NA	NR	NR	NR	50.9	51.6

[a]Mean age and range calculated based on subjects' year of birth; Fitzpatrick grading: I, burns easily, never tans; II, burns easily, tans minimally; III, burns moderately, tans moderately (13).

data for the test product or P2 positive control ($n = 1$); and for technical reasons ($n = 1$). Separate baseline demographics were determined for the FDA final rule cohort and the ISO 24444 cohort. Participants enrolled for the FDA final rule protocol were all white, more than half (69.2%) were female, and the mean age was 41.2 years. Participants had a skin type of Fitzpatrick grade I (30.8%), II (53.8%), or III (15.4%), with a mean ITA of 50.9° (Table II). All participants enrolled for the ISO 24444 protocol were Caucasian, the majority (75.0%) were female, and the mean age was 42.2 years. Participants had a skin type of Fitzpatrick grade I (37.5%), II (50.0%), or III (12.5%), with a mean ITA of 51.6° (Table II).

EFFICACY AND EXPLORATORY ENDPOINTS

Barrier function and moisturizing abilities study. **Dry-skin leg model.** Statistically significant changes in TEWL at 1, 3, and 6 h post–lip balm application indicated an improvement in barrier function (decreased water loss) with test product B and all four comparator/reference products *versus* untreated dry leg skin ($p < 0.0005$) (Table III). Test product B also showed a statistically significant reduction in TEWL, indicating an improvement in barrier function, *versus* three of the four comparator lip balms at all three time points (comparators A, B, and D; $p < 0.0001$); the fourth comparator (comparator C) showed significantly lower TEWL compared with test product B at all three time points ($p < 0.0004$; Table III).

Skicon data indicated a statistically significant improvement in skin hydration (moisturizing ability) for test product B and three of the four comparator lip balms (A, C, and D) *versus* untreated dry leg skin ($p < 0.0001$) at all three time-points (Table III). Improvement in skin hydration was significantly greater for test product B relative to comparators

Table III
Study RH02116: Dry-Skin Leg Model; Barrier Function, and Skin Hydration Data for Test Lip Balm, Four Comparator Products, and Untreated Skin

Time point	Test product B	Comparator A	Comparator B	Comparator C	Comparator D	Untreated skin (control)
Means for TEWL measurements (g m^{-2} h)						
Baseline	5.72	5.83	5.81	5.69	5.77	5.62
1 h	3.40	4.51	4.31	2.61	4.32	5.13
3 h	3.42	4.44	4.36	2.94	4.29	4.92
6 h	3.30	4.32	4.18	2.79	4.16	4.78
Means for Skicon measurements (μS)						
Baseline	38.21	36.82	34.97	35.29	35.46	35.23
1 h	129.37	196.44	52.72	150.68	133.08	41.43
3 h	127.93	179.89	51.52	142.81	118.08	44.14
6 h	131.57	159.03	51.23	126.94	103.75	47.93
Means for Corneometer measurements (au)						
Baseline	12.88	13.56	13.41	12.55	12.89	12.65
1 h	30.99	23.53	16.29	20.33	15.24	14.20
3 h	31.83	24.53	18.04	21.86	16.99	14.60
6 h	31.62	26.89	20.11	25.34	19.77	15.46

μS: micro-siemens.

A and B at 1 h ($p < 0.0001$), 3 h ($p < 0.0001$), and 6 h ($p = 0.001$ and $p < 0.0001$, respectively) post–lip balm application. In addition, test product B demonstrated statistically significant improvement in skin hydration compared with comparator D at 6 h post–lip balm application ($p = 0.0064$; Table III).

The Corneometer data indicated a statistically significant improvement in skin surface hydration for test product B compared with all four comparator/reference lip balms ($p \leq 0.0009$) and untreated dry leg skin ($p < 0.0001$) at all time points postapplication (Table III).

Home-use lip hydration phase. Barrier function and skin hydration data for the four participants who used test product B during the home-use phase of the study are shown in Figure 1.

Mean ± SD TEWL at baseline was 54.60 ± 17.34 g m^{-2} h and decreased to 43.58 ± 4.61 g m^{-2} h after 12 d of test product B application, suggesting a slightly reduced level of water loss. The lowest TEWL score occurred at day 8 (mean ± SD 15.34 ± 4.74 g m^{-2} h). TEWL scores were consistent between the four participants at each time point.

Mean ± SD DMM at baseline was 61.00 ± 13.25 arbitrary units (au) and increased to 63.90 ± 27.62 au by day 12, with a maximum score of 88.30 ± 82.34 au on day 8. This small increase suggests that there was no overall change in skin hydration levels. DMM scores were consistent between three of the participants; the exception was one participant for whom the DMM score was 210.60 au on day 8.

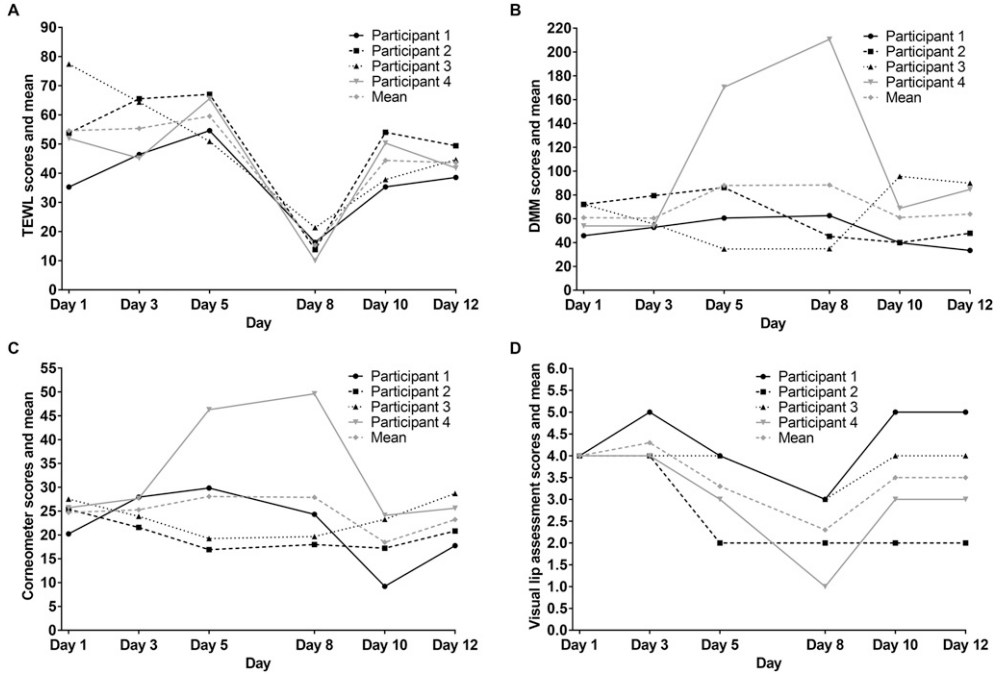

Figure 1. Study RH02116: Home-use lip-hydration phase; barrier function and skin hydration data for four participants over the 11-d study period. Scores and mean values for (A) TEWL, (B) DMM, (C) Corneometer, and (D) visual P2 assessment.

Mean ± SD Corneometer at baseline was 24.72 ± 3.14 au and decreased slightly by day 12 to 23.22 ± 4.87 au, suggesting no overall change in skin hydration levels. There was a high level of variability in the Corneometer scores for all four participants at all time points.

The mean baseline score following visual assessment of lip dryness for the four participants was 4.0. This decreased only marginally to 3.5 by day 12, suggesting no overall impact of the test product on lip dryness. The lowest mean visual score was recorded on day 8 (2.3).

SPF determination studies. For all four SPF studies, the SPF standard P2 positive control achieved a mean SPF within the SD range of the intended SPF according to ISO 24444 (16.4 ± 2.4) and FDA final rule (16.3 ± 3.43), thus confirming the accuracy of the study protocol. The mean value of the tested SPF lip balm formulations did not achieve the intended labeled SPF 15 (Table IV).

In study RH01927, there was a single protocol deviation where one participant underwent initial MED evaluation 18 min outside of the permitted evaluation window. However, this did not affect the study results. Study RH02117 did not achieve the number of valid test results required to determine an SPF value for the test products; for test product B, only nine acceptable results were obtained, and for test product C, only eight acceptable results were obtained.

SAFETY ASSESSMENTS

There were no test product-related AEs or SAEs for any of the five lip balm formulations assessed across the five studies. Furthermore, no local skin reactions, such as irritation, due to product application were observed.

Table IV
Mean SPF Values and Label for Five New Lip Balm Formulations across Four Studies, as per Protocols Defined by the FDA Monograph Final Rule (5) or ISO 24444 (6)

Study	Product identification	Regulatory criteria used	*n* (valid cases)	Mean SPF ± SD	Mean SPF ± CI, %	SPF value and label (protection offered)
RH01927	Test product A	FDA final rule	10	13.7 ± 4.8		10 (low)
RH01927	P2 standard	FDA final rule	10	14.5 ± 2.3		NA[b]
RH01928	Test product A	ISO 24444	13	15.2 ± 3.9		Low
RH01928	P2 standard	ISO 24444	13	14.8 ± 3.2		NA[b]
RH02117	Test product B	FDA final rule	8	13.4 ± 2.4		Invalid data[a]
RH02117	Test product C	FDA final rule	9	13.4 ± 1.5		Invalid data[a]
RH02117	P2 standard	FDA final rule	13	15.4 ± 3.7		NA[b]
RH02385	Test product D	ISO 24444	13		13.9 ± 13.9	10 (low)
RH02385	P2 standard	ISO 24444	13		14.2 ± 12.0	NA[b]
RH02385	Test product E	ISO 24444	12		12.8 ± 16.9	10 (low)
RH02385	P2 standard	ISO 24444	12		13.9 ± 9.9	NA[b]
RH02385	Test product D	FDA final rule	10		14.4 ± 12.6	12 (low)
RH02385	P2 standard	FDA final rule	10		14.4 ± 12.5	NA[b]
RH02385	Test product E	FDA final rule	10		15.5 ± 11.4	11 (low)
RH02385	P2 standard	FDA final rule	10		14.6 ± 12.9	NA[b]

[a]<10 Subjects achieved valid test results, insufficient to determine an SPF value for test products.
[b]Label value not required for standard only test product.

In study RH02116, three AEs were reported but each was considered by the investigator to be unrelated to study treatment. One AE was also a SAE (acute cholecystitis) and occurred during phase 1 of the study. The remaining two AEs (dysmenorrhea and upper respiratory tract infection) were both mild in severity and occurred during the home-use phase of the study.

No AEs were reported in studies RH01928, RH02117, and RH02385. One AE (tingling sensation at the irradiation site of MED #1) was reported in study RH01927 but was not considered to be product related by the investigator because it occurred before product application. This participant continued in the study.

There were no cases of prior or concomitant medication use, or concomitant nondrug treatment/procedures that were considered to have affected the study outcomes.

DISCUSSION

The dry-skin leg model used to help assess lip balm formulations in study RH02116 is a recognized industry method used for the evaluation of moisturizing products on body skin (11,12). Using this model, test product B demonstrated a statistically significant improvement in barrier function and skin hydration across all postapplication time points compared with untreated dry leg skin. Furthermore, test product B significantly improved skin surface hydration compared with four commercially available comparator products, and improved barrier function to a greater extent than three of the comparators. However, although these data are encouraging, they are based on a dry leg skin model. Lip skin differs from skin on the rest of the body in several ways, including the absence of sweat glands (and often sebaceous glands) and by having a thinner SC (18,19). As a result, our data derived from the dry leg skin model will need to be replicated using a dry lip model to determine the specific benefits of test product B for lip skin.

Data from the exploratory novel dry lip model, derived from four different methodologies, produced variable results across all time points. This may be because of the small number of participants that were studied. In addition, the use of multiple methods in parallel to measure skin hydration may have made it more difficult to determine which data were the most relevant and reliable. As an example, the Corneometer method can be less sensitive to immediate changes in hydration of the SC; a method that is sensitive to immediate changes in hydration of the SC is essential when evaluating lip balm formulations. Therefore, further investigation of current methodology is required.

The four SPF determination studies were conducted in accordance with the requirements of FDA final rule (5) and the international standard ISO 24444 (6). The mean SPF for the P2 positive control for all four studies was within the accepted range to meet the relevant regulatory criteria supporting the accuracy of the study protocols. However, although the early stage lip balm formulations tested demonstrated some SPF capability, results from the four SPF studies indicated that each of the formulations failed to meet the requirements for the intended labeled SPF. Consequently, reformulation and retesting will be required. Nevertheless, and importantly, no product-related AEs or SAEs were observed for any of the lip balm formulations assessed.

In study RH01927, test product A achieved a mean SPF value of 13.7; the SPF of this lip balm formulation would be labelled as 10 (low), according to FDA final rule (5), which is lower than the intended SPF 15. In study RH01928, test product A achieved a mean

SPF value of 15.2 according to ISO 24444 criteria (6). In study RH02117, the number of valid SPF results obtained was deemed insufficient, and it was therefore not possible to determine the SPF label for either test product B or test product C. In study RH02385, the mean SPF value for each of test product D and test product E was lower than the intended SPF 15. As a result, both lip balm formulations would have been labelled as low sun protection under the ISO 24444 criteria, and labelled SPF 12 and SPF 11, respectively, under the FDA criteria.

A key strength of the SPF studies that we conducted is that the methodology we used to determine SPF follows industry standards and protocols as defined by the relevant regulators [FDA monograph final rule and international standard ISO 24444 (5,6)].

CONCLUSION

A new range of lip balm formulations was developed to provide sun protection and moisturizing abilities.

A dry-skin leg model was used to assess the potential moisturizing abilities of one of the new lip balm formulations. Using this model, the test product demonstrated improved leg skin hydration and barrier function properties compared with untreated dry skin, and compared with commercially available comparator products. Further testing is required to confirm these benefits on dry lip skin. Data from an exploratory study that used a novel dry-lip hydration model, combining four different assessment methodologies, proved unreliable. Therefore, further investigation of the lip-hydrating potential of the lip balm formulation is needed.

The lip balm formulations studied herein were all rigorously tested to determine their SPF label according to protocols defined by relevant regulatory authorities [FDA monograph final rule and international standard ISO 24444 (5,6)]. The SPF values of the tested lip balm formulations did not achieve the intended SPF levels. As such, they will undergo reformulation and retesting to satisfy the required SPF levels.

ACKNOWLEDGMENTS

The authors would like to thank the participants in these five studies. Editorial support in the form of the development of a draft manuscript in consultation with the authors, and preparation of tables and figures was provided by Victoria Pugh, PhD, of Gardiner-Caldwell Communications, Macclesfield, UK.

FUNDING

These studies were funded by GSK Consumer Healthcare (GSK studies RH01927, RH01928, RH02116, RH02117, and RH02385).

DISCLOSURES

C. Gfeller, G. Hardie, G. Shanga, and H. Mahalingam are all employees of GSK Consumer Healthcare. No additional disclosures to declare.

REFERENCES

(1) B. Dale Wilson, S. Moon, and F. Armstrong, Comprehensive review of ultraviolet radiation and the current status on sunscreens, *J. Clin. Aesthet. Dermatol.*, **5**, 18–23 (2012).

(2) World Health Organization Website, Health Effects of UV Radiation, accessed September 6, 2016, http://www.who.int/uv/health/en/.

(3) W. Stebbins and C.W. Hanke, Lip Cancer: Not Uncommon, Often Overlooked. Skin Cancer Foundation Website, accessed September 6, 2016, http://www.skincancer.org/skin-cancer-information/lip-cancer-not-uncommon.

(4) M.S. Latha, J. Martis, V. Shobha, R. Sham Shinde, S. Bangera, B. Krishnankutty, S. Bellary, S. Varughese, P. Rao, and B. R. Naveen Kumar, Sunscreening agents: a review. *J. Clin. Aesthet. Dermatol.*, **6**, 16–26 (2013).

(5) Department of Health and Human Services, Food and Drug Administration, Sunscreen drug products for over-the-counter human use; final rules and proposed rules, *Fed Regist*, **76**(117) (2011); Part IV: 21 CFR Parts 201, 310, and 352.

(6) ISO 24444:2010(en), Cosmetics—Sun Protection Test Methods—*In vivo* Determination of the Sun Protection Factor (SPF). ISO Website (2010), accessed September 6, 2016, www.iso.org/obp/ui/#iso:std:iso:24444:ed-1:v1:en.

(7) J. Pinnagoda, R. A. Tupker, T. Agner, and J. Serup, Guidelines for transepidermal water loss (TEWL) measurement. A report from the standardization group of the European society of contact dermatitis, *Contact Dermatitis*, **22**, 164–178 (1990).

(8) E. Proksch, J. M. Brandner, and J. M. Jensen, The skin: an indispensable barrier, *Exp. Dermatol.*, **17**, 1063–1072 (2008).

(9) D. Antonov, S. Schliemann, and P. Elsner, Methods for the assessment of barrier function, *Curr. Probl. Dermatol.*, **49**, 61–70 (2016).

(10) M. Obata and H. Tagami, A rapid *in vitro* test to assess skin moisturizers. *J. Soc. Cosmet. Chem.*, **41**, 235–241 (1990).

(11) J. Serup, EEMCO guidance for the assessment of dry skin (xerosis) and ichthyosis: clinical scoring systems. *Skin Res. Technol.*, **1**, 109–114 (1995).

(12) S. L. Zhang, "Monitoring skin hydration by near-infrared spectroscopy and multispectral imaging," in Bioengineering of the Skin: Skin Imaging and Analysis, K. P. Wilhelm, P. Elsner, E. Berardesca, and H. I. Maibach. Eds. (CRC Press, London, 2006).

(13) T. B. Fitzpatrick, The validity and practicality of sun-reactive skin types I through VI. *Arch. Dermatol.*, **124**, 869–871 (1988).

(14) T. B. Fitzpatrick, M. Pathak, and J. A. Parrish, "Protection of human skin against the effects of the sunburn ultraviolet (290–320 nm)," in Sunlight and Man: Normal and Abnormal Photobiologic Responses, M. A. Pathak, L. C. Harber, M. Seiji, T. B. Fitzpatrick, and consulting. Eds. (University of Tokyo Press, Tokyo, 1974), pp. 751–765.

(15) A Cosmetics Europe, Guideline for the Colorimetric Determination of Skin Colour Typing and Prediction of the Minimal Erythemal Dose (med) Without UV Exposure. Cosmetics Europe, Brussels, 2007, accessed September 6, 2016, https://www.cosmeticseurope.eu/publications-cosmetics-europe-association/guidelines.html?view=item&id=63.

(16) J. Ishikawa, H. Yoshida, S. Ito, A. Naoe, T. Fujimura, T. Kitahara, Y. Takema, C. Zerweck, and G. L. Grove, Dry skin in the winter is related to the ceramide profile in the stratum corneum and can be improved by treatment with a eucalyptus extract, *J. Cosmet. Dermatol.*, **12**, 3–11 (2013).

(17) J. Tan, H. Liu, J. J. Leyden, and M. J. Leoni, Reliability of clinician erythema assessment grading scale, *J. Am. Acad. Dermatol.*, **71**, 760–763 (2014).

(18) Z. Ya-Xian, T. Suetake, and H. Tagami, Number of cell layers of the stratum corneum in normal skin—relationship to the anatomical location on the body, age, sex and physical parameters, *Arch. Dermatol. Res.*, **291**, 555–559 (1999).

(19) H. Kobayashi and H. Tagami, Functional properties of the surface of the vermilion border of the lips are distinct from those of the facial skin, *Br. J. Dermatol.*, **150**, 563–567 (2004).