Effects of *Bifidobacterium animalis* subsp. *lactis* Bl-04 on Skin Wrinkles and Dryness: A Randomized, Triple-Blinded, Placebo-Controlled Clinical Trial

Laura Huuskonen 1,*+, Anna Lyra 1, Eunju Lee 2*, Jahyun Ryu 2, Hyunjin Jeong 2, Jihwoon Baek 2, Youngkyoung Seo 2, Minkyung Shin 2,3, Kirsti Tiihonen 1, Tommi Pesonen 4, Antti Lauferma 5, Jenni Reimari 1, Alvin Ibarra 1 and Heli Anglenius 1*

1 IFF Health and Biosciences, Sokeritehtaantie 20, 02460 Kantvik, Finland; anna.l.lyra@iff.com (A.L.); kirsti.tiihonen@iff.com (K.T.); jenni.reimari@iff.com (J.R.); alvin.ibarra@iff.com (A.I.); heli.anglenius@iff.com (H.A.)
2 DERMAPRO Ltd., 61 Seocho-daero, Seocho-gu, Seoul 06570, Korea; seaskyjh@gmail.com (J.R.); meso3550@naver.com (H.J.); jhun100@nate.com (J.B.); youngkyoungseo@gmail.com (Y.S.); haddal@hanmail.net (M.S.)
3 Department of Dermatology, College of Medicine, Kyung Hee University, Kyung Hee Dae Ro 23, Dongdaemun-gu, Seoul 02447, Korea
4 Oy 4Pharma Ltd., Tykistökatu 4D, 20520 Turku, Finland; tommi.pesonen@4pharma.com
5 Skin and Allergy Hospital, Helsinki University Hospital, 00250 Helsinki, Finland; antti.lauferma@hus.fi

*Correspondence: laura.huuskonen@iff.com*

Abstract: The effects of orally consumed probiotics on skin wrinkles and dryness are not fully known. A randomized, placebo-controlled, triple-blinded study was conducted with 148 healthy Korean female volunteers aged between 33 and 60 years, who were administered 1.75×10^9 colony-forming units (CFUs) of *Bifidobacterium animalis* subsp. *lactis* Bl-04 (Bl-04) (N = 74) or matching placebo (N = 74) for 12 weeks. Facial wrinkles (with 3-dimensional (3D) imaging), skin hydration, transepidermal water loss (TEWL), elasticity, and gloss were assessed at baseline and after every 4 weeks of the intervention. Questionnaire-based subjective evaluations of product efficacy and usability were also analyzed. The consumption of Bl-04 was safe and ameliorated significantly facial skin wrinkle parameters (total wrinkle area and volume, average depth of wrinkles, and arithmetic average roughness (Ra)) versus placebo at 4 weeks, but there were no differences at Week 8 or 12 between groups. Skin hydration, TEWL, elasticity, and gloss were similar between treatment groups, as were the subjective evaluation scores. Oral consumption of Bl-04 indicated promising short-term effects on skin appearance from the winter toward the spring. In future study designs, special attention should be paid to environmental conditions as well as to the skin condition and age of the participants.

Keywords: *Bifidobacterium lactis*; facial wrinkles; skin hydration; TEWL; weather; probiotic; randomized clinical trial

1. Introduction

Healthy skin is a barrier between the internal and external environment, and its chief function is to protect the body from external insults and prevent water from evaporating from the inside [1,2]. The dermis and multilayered epidermis are important for maintaining the integrity of this barrier [3]. The stratum corneum of the epidermis is the primary layer that mediates the permeability of the skin barrier; the dermis, which lies beneath the stratum corneum, comprises connective tissue that gives the skin its elasticity, strength, and ability to resist external forces [3,4].

Skin is metabolically active and regenerates throughout the human lifespan, but the capacity for the latter declines with age [2]. Various external and internal factors affect the appearance of the skin and the rate at which it ages [5]. Skin aging can be divided into
chronological aging, caused by biological factors, and photoaging, which evolves due to exposure to ultraviolet (UV) radiation; both subtypes manifest as dryness, loss of elasticity, wrinkling, laxity, reduced gloss, and increased pigmentation [4].

Dietary factors and balanced nutrition have been suggested to influence skin aging [4]. The gut–skin axis is a loosely defined term that describes the interplay between the gut, immune system, and the skin. Diseases such as inflammatory bowel disease, celiac disease, psoriasis, acne, and atopic dermatitis (AD) have been shown to have etiological links to the gut and skin, confirming the existence of the gut–skin axis [6,7].

Probiotics are defined as “live microorganisms that, when administered in adequate amounts, confer a health benefit on the host” [8]. Probiotics have provided benefits in the treatment and prevention of gastrointestinal diseases and have several mechanisms for modulating the mucosal immune system, which has been implicated as the key element in mediating these health benefits [9,10]. Because certain probiotics function through maintenance of the gut’s barrier and immune system, they have been suggested as one means of preserving healthy skin via the gut–skin axis [6]. Evidence for the efficacy of orally consumed probiotics in preserving healthy skin has been found in studies with regard to various outcomes, including skin hydration, transepidermal water loss (TEWL), elasticity, or wrinkles [11–13], hydration in capsaicin-induced sensitivity assay [13], self-reported skin roughness and dryness [13], and cutaneous immune system after UV exposure [14].

Bifidobacterium animalis ssp. lactis Bl-04 (Bl-04) has long been considered to be safe and suitable for human consumption [15]. It can tolerate gastrointestinal conditions in vitro [16–18] and remains viable when consumed [19,20]. The immunological effects of Bl-04 have gained particular interest. The ability of Bl-04 to modulate certain immune functions in humans has been assessed in a vaccination study and in an experimental rhinovirus challenge model [21–24].

The probiotic effects of Bl-04 have thus been evaluated in several clinical studies, but its effects on the skin have remained unexamined. The purpose of this randomized, triple-blinded, placebo-controlled, parallel-group clinical trial was to explore the effects of orally consumed Bl-04 on two primary objectives—skin wrinkles and hydration—and several secondary objectives: TEWL, elasticity, gloss, and subjective questionnaires on product efficacy and usability, among females participants with facial wrinkles and dry skin, and aged between 30 and 60 years per their Korean age (corresponding to 29 to 59 years by international age; at birth a Korean is already 1 year old, turning 1 year older on January 1).

2. Materials and Methods
2.1. Trial Design
This randomized, triple-blind, placebo-controlled, parallel-group clinical trial was performed at Dermapro Ltd. Skin Research Center (Seoul, Korea) according to the Good Clinical Practice (GCP) guidelines (Integrated addendum to ICH E6(R1): Guideline for Good Clinical Practice) E6(R2) and the WMA Declaration of Helsinki (64th WMA General Assembly, Fortaleza, Brazil, October 2013).

The study was conducted between November 2020 and April 2021 and comprised 5 visits: a screening visit (V1); a baseline visit, during which the participants were randomized (V2); and 3 follow-up visits (V3 on Day 28 ± 2, V4 on Day 56 ± 2, and V5 on Day 84 ± 2) during the intervention (Figure 1). At V1, all participants were informed about the content and purpose of the study and signed written consent forms to participate in the trial.

Eligibility criteria were checked at V1, including an assessment of facial skin wrinkles by visual scale, skin hydration of the cheek with a Corneometer®, and a blood safety test. Demographic data, relevant medical history, and concomitant medications (CMs) were also collected.

Skin care products devoid of probiotics were distributed to the participants for use during the study, which the subjects were advised to use from 3 weeks before V2 (from Day-21) until the end of the study. They were also advised to maintain their lifestyle
Eligibility criteria were checked at V1, including an assessment of facial skin wrinkles \( (R_v) \) (representing facial wrinkles) in the treatment group and a standard deviation of 15 \( \mu m \).

Outcome measures for facial wrinkles, hydration, TEWL, elasticity, gloss, and facial photographing were performed at V2–V5. Subjective questionnaires on skin improvement were completed at V3, V4, and V5, and product usability at V5. Blood safety tests were performed at V5 (in addition to V1), and body composition was analyzed at baseline and V5.

Figure 1. Study outline. Visit 1 (V1) was a screening visit at which all potential participants were informed about the content of the study, written consent forms to participate in the study were signed, and eligibility was confirmed. The baseline visit (Visit 2; V2) and all follow-up visits (Visits 3 to 5; V3 to V5) included assessments of skin wrinkles, hydration, transdermal water loss (TEWL), elasticity, gloss, facial photography, and a questionnaire on the perceived effects on the skin (only on V3 to V5). Questionnaire on product usability was filled on V5. Compliance with the study product, adverse events (AEs), and concomitant medication (CM) use were also recorded.

2.2. Ethical Approval and Registry

The ethical and scientific validity of this study was reviewed by the Dermapro Ltd. institutional review board (IRB) (Seoul, Korea; Approval 1-220777-A-N-01-DICN20181 dated 6 August 2020), and the study was performed with the voluntary consent of the participants. The trial was registered with ISRCTN (ISRCTN10924200) and CRIS (KCT0005565) before its outset.

2.3. Sample Size

The sample size was calculated by a two-group t-test with a two-sided significance level of 5%. Calculations were made separately for both primary outcomes, skin wrinkles, and hydration. The estimations were based on studies with comparable study designs, populations, and endpoint definitions ([11] for skin wrinkles and [11,25–30] for skin hydration).

With 80% power to detect a difference with 95% confidence, 64 evaluable participants per group were necessary to observe a difference of 7.5 \( \mu m \) in maximum profile valley depth (Rv) (representing facial wrinkles) in the treatment group and a standard deviation of 15 \( \mu m \).

For skin hydration, to detect a 5-unit difference with a standard deviation of 8 units in the treatment group, 42 evaluable participants per group were required, with 80% power and 95% confidence.

Thus, considering an attrition rate of 30%, a total of 170 randomized participants were to be allocated to the study to ultimately obtain 128 evaluable participants.

2.4. Inclusion and Exclusion Criteria

A total of 170 healthy Korean female volunteers were screened for entrance into the study. The inclusion and exclusion criteria were as follow:

Inclusion criteria:

- Healthy female volunteers aged 20-65 years.
- No history of skin diseases or disorders.
- No use of topical skin care products within the past 4 weeks.
- No history of allergy or sensitivity to the study products.

Exclusion criteria:

- Known history of severe skin conditions such as acne, eczema, or psoriasis.
- Use of systemic or topical medications that may affect skin.
- Pregnancy or lactation.
- Participation in any other clinical trials within the past 4 weeks.
- Allergy or sensitivity to study products.

Informed consent was obtained from all participants before enrollment in the study.
1. Korean female subjects aged 30 to 60 years (by Korean age).
2. Dry skin on cheek (hydration value < 48 arbitrary units (AU) on a Corneometer®).
3. Greater than grade 3 skin wrinkle per a DERMAPRO standard photograph.
4. No chronic or acute disease, including skin disease.
5. Signed informed consent for study participation.
6. Cooperation and availability for follow-up during the study period.

Exclusion criteria
1. Consumption of probiotics as dietary supplements, food, or beverage during the previous 2 weeks.
2. Pregnancy, planned pregnancy, or lactation.
3. Irritation or symptomatic allergy to food, including ingredients of cosmetic, medical, and test products.
4. Use of oral or topical antibiotics during the previous 3 months.
5. Use of oral retinoid/steroid drug or topical steroid application during the previous 6 months.
6. Use of functional cosmetics to improve skin wrinkles, hydration, or elasticity within the past 3 months.
7. Having had skin treatment on the test site (e.g., decortication and Botox treatment).
8. Participation in a previous study without an appropriate intervening period (3 months) between studies.
9. Presence of disease that could affect the aim of the study (e.g., cardiovascular, kidney, liver, thyroid, gastrointestinal disease, gout).
10. Presence of any skin disease (e.g., AD) on test site.
11. Presence of any chronic disease (e.g., diabetes, asthma, high blood pressure) or psychiatric disorder (e.g., depression, schizophrenia).
12. Use of medication for obesity (e.g., antidepressants, anorectics), contraceptives, hormones, or diuretics.
13. Excessive alcohol use (over 30 g alcohol per day) or drug abuse.
14. Sensitive or hypersensitive skin.
15. Damaged skin in or around the test area (including sunburn, tattoos, scars, or other disfiguration of the test area).
16. Abnormal clinical chemical analysis result at V1, per a medical specialist.
17. Presence of a problem that could interfere with the aim of the study, based on the judgment of the principal investigator.

2.5. Interventions
Participants received one daily capsule of 1.75 × 10^9 colony-forming units (CFUs) BI-04 (ATCC SD5219) in microcrystalline cellulose with magnesium stearate and silicon dioxide as flow agents as a probiotic supplement or a comparable capsule without BI-04 as a placebo for 12 weeks. Participants were instructed to consume one capsule orally before breakfast on the following morning after V2 to the morning of V5. The IP was manufactured by Danisco USA Inc. (Madison, WI, USA).

2.6. Randomization and Blinding
Randomization to the BI-04 and placebo arms was performed at a 1:1 ratio using nQuery Advisor 7.0, version 7.0.1490.0 (San Diego, CA, USA), applying a random seed and a block size of 4 units. No stratification was conducted. In the randomization, eligible participants were enrolled sequentially by the study investigator at V2 and assigned the next free randomization number that was indicated on the IP label. Participants who dropped out of the study following the randomization were not replaced.

The study was conducted and analyzed in a triple-blind manner (i.e., the investigator, participants, and statistician were blinded). The randomization list (randomization numbers 001 to 200) was created by the IP manager, who was responsible for packaging and labeling the IP and creating individual emergency unblinding envelopes at Danisco USA.
Inc. (Madison, WI, USA; the IP manager did not otherwise participate in the study). The intactness of the emergency unblinding envelopes was monitored over the course of the study. After the database lock and blind data review, the randomization code was opened as Groups A and B for the statistical analyses without revealing the allocation. After the analyses were completed and the clinical study report was signed, the treatment codes were revealed as Bl-04 and placebo.

2.7. Study Assessments

All skin assessments were performed after 30 min of stabilization at a controlled temperature of 22 ± 2 °C and a relative humidity of 50 ± 5%. In addition, before each study visit, participants were advised not to have applied makeup for 12 h.

2.7.1. Primary Outcomes: Facial Wrinkles and Skin Hydration

Skin wrinkles were measured from the corner of the eye at V2–V5 using the PRIMOS® Premium 3-dimensional (3D) imaging system, software version 5.8 (GFMesstechnik GmbH, Teltow, Germany). Nine parameters were analyzed automatically from images that were taken from the same area of the corner of the eye at each visit: average depth of wrinkles, mean depth of the largest wrinkle, maximum depth of the largest wrinkle, total wrinkle area, total wrinkle volume, the total length of wrinkles, arithmetic mean deviation of the profile (Ra), the maximum height of the profile (Ry), and 10-point height of irregularities (Rz). To analyze the exact same location, the images that were taken during subsequent visits were aligned to the baseline image using the device software.

Skin hydration was measured with a Corneometer® CM 825 (Courage & Khazaka, Cologne, Germany) and a Moisturemeter® D (Delfin, Kuopio, Finland) on the cheek (crossing area between the center of the pupil and the tip of the nose), the middle of the forearm (10 cm above the wrist, avoiding visible blood vessels), and the back of the hand (crossing area between the middle finger and wrist, avoiding bone and visible blood vessels). The S15 probe was used with the Moisturemeter® D to measure skin hydration up to a depth of 1.5 mm. With both devices, each location was measured in triplicate, and the average value was used in the analysis.

2.7.2. Secondary Outcomes

TEWL was measured from the cheek, middle of the forearm, and back of the hand (the same locations as the measurements for skin hydration) with a Tewameter® TM300 (Courage & Khazaka). These measurements were repeated within 30 s, and the average of the last 3 stabilized values was used in the analysis.

Skin elasticity was recorded with a Cutometer® MPA580 (Courage & Khazaka) at the cheek (crossing area between the center of the pupil and the tip of the nose). The measurement was performed with a 2-s on/off pulse of a constant 450-mbar negative pressure, drawing the skin into the aperture of the probe, followed by its release. Three cycles of suction and release constituted one measurement cycle. The parameter R2 (gross elasticity, Ua/Uf) was used for the primary analysis.

Skin gloss was measured with a SkinGlossMeter® (Delfin) on the cheek (crossing area between the center of the pupil and the tip of the nose). The measurement was repeated three times, and the average value was used for analysis.

Facial images were taken with a VISIA® CR (Canfield Scientific, Parsippany, NJ, USA) with standard 2 and parallel-polarized setting.

Participants answered the subjective evaluation questionnaires on product efficacy and usability using a 5-point Likert scale: 1. Difficult/Dissatisfied/Disagree; 2. Somewhat difficult/dissatisfied/disagree; 3. No opinion; 4. Somewhat easy/satisfied/agree; and 5. Easy/Satisfied/Agree. The frequency and percentage of positive answers (Responses 4 and 5) were analyzed for each questionnaire.

Body composition parameters—weight, fat mass, body mass index, body fat percentage, and waist–hip ratio—were measured with an Inbody® 270 (Biospace, Seoul, Korea).
2.7.3. Safety Outcomes and Other Evaluations

Blood glucose, hemoglobin, serum glutamic-oxaloacetic transaminase (SGOT), serum glutamic-pyruvic transaminase (SGPT), gamma-glutamyl transpeptidase ($\gamma$-GTP), total cholesterol (T-cholesterol), triglycerides, high-density lipoprotein (HDL-cholesterol), low-density lipoprotein (LDL-cholesterol), and creatinine were analyzed with regard to safety at V1 and V5 (Daehang Hospital, Seoul, Korea).

AEs were recorded based on clinical observation by the investigator and as reported by the participants. The frequency, duration, and intensity of the AEs and their relationship with the test product were estimated by the investigator.

Product compliance was calculated, based on the number of leftover capsules in the container that was returned to the clinic.

In an ancillary evaluation, outdoor temperature (mean, min, max), mean relative humidity, and mean UV radiation (Korea Meteorological Administration, Seoul, Korea) on weekdays between the first participant's initial visit to the last participant's final visit were recorded.

Participant age, medical history, relevant CM use, skin, and lifestyle characteristics were collected as demographic data. The participants evaluated the following aspects of skin and lifestyle: skin type on the face and body (dry, normal, oily, dry and oily, or problematic with acne or other skin disease), skin hydration on the face and body (sufficient, normal, deficient), skin sebum on the face and body (glossy, normal, deficient), skin surface roughness on the face and body (smooth, normal, rough), facial skin thickness (thin, normal, thick), average daily time of UV exposure (less than 1 h, 1–3 h, over 3 h), average hours of sleep per day (less than 5, 5–8, over 8), the average amount of daily smoking (none, fewer than 10 pieces, over 10 pieces), the sensitivity of skin (yes, no), development of stinging sensation within 30 min of applying cosmetics (yes, no), and having experienced adverse reactions from cosmetics in the past 12 months? (yes, no).

2.8. Statistical Methods

The statistical analyses were conducted using SPSS® (IBM, Chicago, IL, USA). All data were analyzed for intention-to-treat (ITT) and per-protocol (PP) populations; the primary population was PP. Participants who had missing endpoint data, had a compliance rate of less than 80% IP use, used prohibited CMs, or encountered other major deviations that could have affected the primary endpoints were excluded from the PP population. Normality was determined by Kurtosis and Skewness test, and for the between-group analyses, homogeneity at baseline was tested by independent t-test.

For the within-group comparisons between time points (V2 vs. V3, V4, or V5), repeated measures analysis of variance (RM-ANOVA) was applied for parametric variables, and the Friedman test and Wilcoxon Signed-rank test with Bonferroni correction were used for nonparametric variables. For between-group comparisons, RM-ANOVA was performed for parametric values, and Mann–Whitney U-test was used for nonparametric values.

Clinical blood tests and body composition parameters from V1 and V2, respectively, were compared with V5 by paired t-test.

The subjective questionnaires on IP efficacy and usability were analyzed by chi-square test and Fisher’s Exact test.

This study was an exploratory study in nature, and no adjustments for multiplicity were performed. A statistically significant difference was set at $p < 0.05$.

2.9. Post Hoc Analyses

The primary study results were examined further by post hoc analysis, evaluating the effects in subgroups with the appropriate participant number, and correlations between a skin variable and weather parameter were evaluated by scatterplot and Pearson correlation. The post hoc analyses were conducted using SAS® for Windows, version 9.4 (SAS Institute Inc., Cary, NC, USA). The post hoc analysis results are shown only for the subgroups by age: 30–44 years and 45–60 years by international age.
Repeated measures analysis of covariance (RM-ANCOVA) was applied for continuous variables to analyze differences between groups and the changes within groups. The models included fixed effects for baseline value, treatment group, repeating factor week (Week 4 (V3), Week 8 (V4), Week 12 (V5)), and the interaction between treatment group and week. ‘Participant’ was the random effect in the model. An unstructured covariance matrix was assumed for the repeated measures for each participant over time. The contrasts between and within groups were estimated from the models, with 95% confidence intervals (CIs), using least square means. Two-sided \( p \)-values were calculated from the model. Further modeling was performed, including an adjustment factor as an additional covariate in the model.

Model fit was evaluated by testing residual normality using the Shapiro–Wilk test. If residual normality was not achieved with standard transformations (e.g., log, square root, inverse), nonparametric methods were applied. In such cases, an overall comparison between treatment groups was performed by Friedman test, within-group differences were analyzed by time point using the Wilcoxon Signed-rank test, and between-group differences were examined by time point using Mann–Whitney U-test.

Correlations between two continuous variables were evaluated using scatter plots and Pearson correlation coefficients.

A statistically significant difference was set at \( p < 0.05 \) and two-sided 95% CIs were produced for the estimated differences. Missing values were not imputed in the analyses.

3. Results

3.1. Recruitment and Study Populations

Volunteers were recruited through advertisements and clinic volunteer registry, followed by prescreening phone calls to schedule screening visits. Of the 170 volunteers who were screened, 148 were randomly allocated into two parallel, equal-sized intervention groups—Bl-04 and placebo—and 132 completed the study. The ITT population (N = 147; 74 active and 73 placebo) comprised all randomized participants who had consumed the IP, except for one participant who was subjected to emergency unblinding during the study. The PP analysis included 128 participants (61 and 67 in the Bl-04 and placebo groups, respectively; Figure 2). The safety population was composed of all randomized participants who had consumed the IP, including the participant who underwent emergency unblinding (Figure 2).

3.2. Baseline Data

A total of 147 females participated in the study, aged between 33 and 60 (49.8 ± 6.0 years (mean ± SD)). In the Bl-04 group, the average age was 49.2 ± 6.1 years (mean ± SD, N = 74), versus 50.4 ± 5.9 years (N = 73) in the placebo group. Skin hydration (by Corneometer®) at V1 was 43.7 ± 3.8 AU (mean ± SD) in the Bl-04 group and 45.0 ± 3.1 AU in the placebo group. Self-evaluated skin and lifestyle characteristics at the screening (Table 1 and Supplementary Material Table S1) and the skin parameters (wrinkle parameters, hydration, TEWL, skin elasticity, and skin gloss) at baseline (Tables 2–4, and Supplementary Material Tables S2–S4) were comparable between groups.

Of the body composition parameters (Supplementary Material Table S6), only body weight differed significantly at baseline between the Bl-04 and placebo groups (\( p = 0.042 \)), which had mean weights of 61.0 kg and 57.7 kg, respectively.

3.3. Primary Outcomes

3.3.1. Skin Wrinkles

The effect of Bl-04 on facial wrinkles in the corner of the eye was analyzed by optical 3D imaging PRIMOS® premium at baseline and Weeks 4, 8, and 12 (Table 2). The mean values at each time point (Table 2) were used to calculate relative changes from baseline (Figure 3a–d).
Figure 2. CONSORT flow diagram. Of the 170 volunteers who were screened for eligibility, 148 were randomly allocated into Bl-04 and placebo and 132 completed the study. The intention-to-treat (ITT) population included only 147 participants due to one participant being subjected to emergency unblinding during the study. According to subject classification criteria, the participants in the per-protocol (PP) population were to have at least an 80% compliance rate with the investigational product (IP), a minimum of 14 days of use of the provided skin care products prior to randomization, less than $-5$ or $+7$ days deviation from the scheduled visits, no missing endpoint data on any visit, under 10% acclimatation condition deviation prior to skin measurements, no treatment of or reaction on the skin test site, no considerable lifestyle change during the study, and no use of prohibited concomitant medications.

Table 1. Self-evaluated baseline skin and lifestyle characteristics in the PP population.

<table>
<thead>
<tr>
<th>Item</th>
<th>Classification</th>
<th>BI-04 (N = 61)</th>
<th>Placebo (N = 67)</th>
<th>Total (N = 128)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Facial skin type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td>35 (57.4)</td>
<td>41 (61.2)</td>
<td>76 (59.4)</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>19 (31.1)</td>
<td>18 (26.9)</td>
<td>37 (28.9)</td>
<td></td>
</tr>
<tr>
<td>Oily</td>
<td>1 (1.6)</td>
<td>1 (1.5)</td>
<td>2 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Dry and oily</td>
<td>6 (9.8)</td>
<td>7 (10.4)</td>
<td>13 (10.2)</td>
<td></td>
</tr>
<tr>
<td>Problematic (acne or other skin disease)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td></td>
</tr>
<tr>
<td><strong>Body skin type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td>35 (57.4)</td>
<td>39 (58.2)</td>
<td>74 (57.8)</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>23 (37.7)</td>
<td>24 (35.8)</td>
<td>47 (36.7)</td>
<td></td>
</tr>
<tr>
<td>Oily</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td></td>
</tr>
<tr>
<td>Dry and oily</td>
<td>3 (4.9)</td>
<td>4 (6.0)</td>
<td>7 (5.5)</td>
<td></td>
</tr>
<tr>
<td>Problematic (acne or other skin disease)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Item Classification</th>
<th>BI-04 (N = 61)</th>
<th>Placebo (N = 67)</th>
<th>Total (N = 128)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%) 1</td>
<td>N (%) 1</td>
<td>N (%) 1</td>
</tr>
<tr>
<td><strong>Facial skin hydration</strong> 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sufficient</td>
<td>1 (1.6)</td>
<td>0 (0.0)</td>
<td>1 (0.8)</td>
</tr>
<tr>
<td>Normal</td>
<td>27 (44.3)</td>
<td>25 (37.3)</td>
<td>52 (40.6)</td>
</tr>
<tr>
<td>Deficient</td>
<td>32 (52.5)</td>
<td>42 (62.7)</td>
<td>74 (57.8)</td>
</tr>
<tr>
<td><strong>Body skin hydration</strong> 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sufficient</td>
<td>0 (0.0)</td>
<td>1 (1.5)</td>
<td>1 (0.8)</td>
</tr>
<tr>
<td>Normal</td>
<td>28 (45.9)</td>
<td>30 (44.8)</td>
<td>58 (45.3)</td>
</tr>
<tr>
<td>Deficient</td>
<td>32 (52.5)</td>
<td>35 (52.2)</td>
<td>67 (52.3)</td>
</tr>
<tr>
<td><strong>Facial skin sebum</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glossy</td>
<td>3 (4.9)</td>
<td>2 (3.0)</td>
<td>5 (3.9)</td>
</tr>
<tr>
<td>Normal</td>
<td>29 (47.5)</td>
<td>37 (55.2)</td>
<td>66 (51.6)</td>
</tr>
<tr>
<td>Deficient</td>
<td>29 (47.5)</td>
<td>28 (41.8)</td>
<td>57 (44.5)</td>
</tr>
<tr>
<td><strong>Body skin sebum</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glossy</td>
<td>1 (1.6)</td>
<td>1 (1.5)</td>
<td>2 (1.6)</td>
</tr>
<tr>
<td>Normal</td>
<td>27 (44.3)</td>
<td>37 (55.2)</td>
<td>64 (50.0)</td>
</tr>
<tr>
<td>Deficient</td>
<td>33 (54.1)</td>
<td>29 (43.3)</td>
<td>62 (48.4)</td>
</tr>
<tr>
<td><strong>Facial skin surface roughness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smooth</td>
<td>9 (14.8)</td>
<td>20 (29.9)</td>
<td>29 (22.7)</td>
</tr>
<tr>
<td>Normal</td>
<td>47 (77.0)</td>
<td>37 (55.2)</td>
<td>84 (65.6)</td>
</tr>
<tr>
<td>Rough</td>
<td>5 (8.2)</td>
<td>10 (14.9)</td>
<td>15 (11.7)</td>
</tr>
<tr>
<td><strong>Body skin surface roughness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smooth</td>
<td>8 (13.1)</td>
<td>13 (19.4)</td>
<td>21 (16.4)</td>
</tr>
<tr>
<td>Normal</td>
<td>43 (70.5)</td>
<td>46 (68.7)</td>
<td>89 (69.5)</td>
</tr>
<tr>
<td>Rough</td>
<td>10 (16.4)</td>
<td>8 (11.9)</td>
<td>18 (14.1)</td>
</tr>
<tr>
<td><strong>Facial skin thickness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thin</td>
<td>17 (27.9)</td>
<td>25 (37.3)</td>
<td>42 (32.8)</td>
</tr>
<tr>
<td>Normal</td>
<td>40 (65.6)</td>
<td>40 (59.7)</td>
<td>80 (62.5)</td>
</tr>
<tr>
<td>Thick</td>
<td>4 (6.6)</td>
<td>2 (3.0)</td>
<td>6 (4.7)</td>
</tr>
<tr>
<td><strong>Average daily UV exposure time</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 1 h</td>
<td>25 (41.0)</td>
<td>32 (47.8)</td>
<td>57 (44.5)</td>
</tr>
<tr>
<td>1-3 h</td>
<td>33 (54.1)</td>
<td>33 (49.3)</td>
<td>66 (51.6)</td>
</tr>
<tr>
<td>More than 3 h</td>
<td>3 (4.9)</td>
<td>2 (3.0)</td>
<td>5 (3.9)</td>
</tr>
</tbody>
</table>

1 N (Frequency) = Number of answers, % (Percentage) = Number of answers/Total number of subjects (61 or 67) x 100. 2 One participant in the BI-04 group was excluded due to supplying an answer that was not in the classification. 3 One participant each in the BI-04 and placebo groups was excluded due to supplying an answer that was not in the classification.

Table 2. Skin wrinkle parameters by optical 3-dimensional (3D) imaging system PRIMOS® premium, from the corner of the eye in the PP population every 4 weeks from baseline to Week 12.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Time Point</th>
<th>BI-04 (N = 61) Mean (SD) 4 p-Value 1</th>
<th>Placebo (N = 67) Mean (SD) 4 p-Value 1</th>
<th>p-Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (SD) 4</td>
<td>p-Value 1</td>
<td></td>
</tr>
<tr>
<td>Average depth of wrinkles, µm</td>
<td>Baseline 37.01 (9.28) - 37.69 (9.71)</td>
<td>0.470</td>
<td>0.470</td>
<td></td>
</tr>
<tr>
<td>Mean depth largest wrinkle, µm</td>
<td>Baseline 52.43 (20.00) - 53.79 (20.34)</td>
<td>0.279</td>
<td>0.279</td>
<td></td>
</tr>
<tr>
<td>Max. depth largest wrinkle, µm</td>
<td>Baseline 152.84 (77.98) - 153.87 (81.68)</td>
<td>0.734</td>
<td>0.734</td>
<td></td>
</tr>
<tr>
<td>Total wrinkle area, mm²</td>
<td>Baseline 42.13 (0.31) - 42.13 (0.28)</td>
<td>0.982</td>
<td>0.982</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2. Cont.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Time Point</th>
<th>BI-04 (N = 61)</th>
<th></th>
<th>Placebo (N = 67)</th>
<th></th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (SD) 4</td>
<td>p-Value 1</td>
<td>Mean (SD) 4</td>
<td>p-Value 1</td>
<td>p-Value 2</td>
</tr>
<tr>
<td>Total wrinkle volume, mm³</td>
<td>Baseline</td>
<td>1.56 (0.40)</td>
<td>-</td>
<td>1.61 (0.36)</td>
<td>-</td>
<td>0.490</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>1.59 (0.42)</td>
<td>0.280</td>
<td>1.71 (0.40)</td>
<td>&lt;0.001 *</td>
<td>0.019 *</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>1.61 (0.42)</td>
<td>0.004 *</td>
<td>1.66 (0.38)</td>
<td>0.057</td>
<td>0.863</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>1.61 (0.40)</td>
<td>0.041 *</td>
<td>1.62 (0.36)</td>
<td>0.445</td>
<td>0.309</td>
</tr>
<tr>
<td>Total length of wrinkles, mm</td>
<td>Baseline</td>
<td>109.23 (16.26)</td>
<td>-</td>
<td>106.00 (11.80)</td>
<td>-</td>
<td>0.205</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>108.46 (15.05)</td>
<td>0.367</td>
<td>105.42 (11.64)</td>
<td>0.441</td>
<td>0.778</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>108.74 (15.22)</td>
<td>0.511</td>
<td>104.99 (12.32)</td>
<td>0.163</td>
<td>0.388</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>108.08 (15.45)</td>
<td>0.203</td>
<td>105.70 (11.36)</td>
<td>0.724</td>
<td>0.787</td>
</tr>
<tr>
<td>Ra, µm</td>
<td>Baseline</td>
<td>19.17 (4.28)</td>
<td>-</td>
<td>19.62 (3.92)</td>
<td>-</td>
<td>0.533</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>19.58 (4.59)</td>
<td>0.138</td>
<td>20.83 (4.41)</td>
<td>&lt;0.001 *</td>
<td>0.031 *</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>19.90 (4.65)</td>
<td>0.001 *</td>
<td>20.21 (4.14)</td>
<td>0.030 *</td>
<td>0.661</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>19.87 (4.49)</td>
<td>0.005 *</td>
<td>19.86 (3.93)</td>
<td>0.294</td>
<td>0.170</td>
</tr>
<tr>
<td>Ry, µm</td>
<td>Baseline</td>
<td>307.84 (109.65)</td>
<td>-</td>
<td>311.52 (91.50)</td>
<td>-</td>
<td>0.836</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>324.39 (116.72)</td>
<td>0.010 *</td>
<td>328.49 (104.48)</td>
<td>0.007 *</td>
<td>0.963</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>316.17 (108.98)</td>
<td>0.654 3</td>
<td>314.40 (93.84)</td>
<td>0.636</td>
<td>0.530</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>310.13 (94.70)</td>
<td>1.372 3</td>
<td>314.02 (95.41)</td>
<td>0.719</td>
<td>0.983</td>
</tr>
<tr>
<td>Rz, µm</td>
<td>Baseline</td>
<td>234.86 (57.64)</td>
<td>-</td>
<td>240.68 (53.42)</td>
<td>-</td>
<td>0.554</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>241.84 (59.58)</td>
<td>0.053</td>
<td>249.61 (55.36)</td>
<td>0.007 *</td>
<td>0.684</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>241.89 (61.75)</td>
<td>0.019 *</td>
<td>242.73 (51.97)</td>
<td>0.497</td>
<td>0.238</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>240.32 (57.54)</td>
<td>0.108</td>
<td>241.23 (50.45)</td>
<td>0.869</td>
<td>0.298</td>
</tr>
</tbody>
</table>

1 Within-group comparisons to baseline by Repeated measures analysis of variance (RM-ANOVA). Significant p-values (p < 0.05) are marked with an asterisk (*). 2 Between-group comparisons at baseline by independent t-test and at Weeks 4, 8, and 12 by RM-ANOVA. Significant p-values (p < 0.05) are marked with an asterisk (*). 3 Wilcoxon Signed-rank test with Bonferroni correction was applied when values were not normally distributed. 4 Values shown as mean ± standard deviation (SD).

### Table 3. Skin hydration values in the PP population, as measured with the Corneometer® and expressed as arbitrary units (AUs). Skin hydration was measured from the cheek, forearm, and hand every 4 weeks from baseline to Week 12.

<table>
<thead>
<tr>
<th>Location</th>
<th>Time Point</th>
<th>BI-04 (N = 61)</th>
<th></th>
<th>Placebo (N = 67)</th>
<th></th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (SD) 4</td>
<td>p-Value 1</td>
<td>Mean (SD) 4</td>
<td>p-Value 1</td>
<td>p-Value 2</td>
</tr>
<tr>
<td>Cheek, AU</td>
<td>Baseline</td>
<td>44.87 (2.68)</td>
<td>-</td>
<td>45.24 (2.01)</td>
<td>-</td>
<td>0.697 4</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>45.47 (6.47)</td>
<td>2.940 3</td>
<td>46.27 (6.16)</td>
<td>1.045 3</td>
<td>0.615 4</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>45.97 (7.96)</td>
<td>2.488 3</td>
<td>46.87 (7.35)</td>
<td>0.587 3</td>
<td>0.580 4</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>46.12 (6.94)</td>
<td>0.797 3</td>
<td>46.67 (7.12)</td>
<td>0.485 3</td>
<td>0.780 4</td>
</tr>
<tr>
<td>Forearm, AU</td>
<td>Baseline</td>
<td>35.63 (7.89)</td>
<td>-</td>
<td>37.40 (8.61)</td>
<td>-</td>
<td>0.227</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>34.51 (7.20)</td>
<td>0.154</td>
<td>35.60 (6.65)</td>
<td>0.071</td>
<td>0.588</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>34.37 (6.78)</td>
<td>0.153</td>
<td>36.01 (7.53)</td>
<td>0.156</td>
<td>0.917</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>33.98 (7.20)</td>
<td>0.061</td>
<td>34.28 (6.83)</td>
<td>0.003 *</td>
<td>0.277</td>
</tr>
<tr>
<td>Hand, AU</td>
<td>Baseline</td>
<td>40.36 (8.39)</td>
<td>-</td>
<td>40.44 (7.41)</td>
<td>-</td>
<td>0.957</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>37.34 (7.24)</td>
<td>&lt;0.001 *</td>
<td>38.41 (7.35)</td>
<td>0.005 *</td>
<td>0.349</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>37.59 (7.23)</td>
<td>0.002 *</td>
<td>38.51 (7.95)</td>
<td>0.019 *</td>
<td>0.473</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>37.69 (6.90)</td>
<td>0.001 *</td>
<td>36.81 (7.69)</td>
<td>&lt;0.001 *</td>
<td>0.386</td>
</tr>
</tbody>
</table>

1 Within-group comparisons to baseline by RM-ANOVA. Significant p-values (p < 0.05) are marked with an asterisk (*). 2 Between-group comparisons at baseline by independent t-test and at Weeks 4, 8, and 12 by RM-ANOVA. Significant p-values (p < 0.05) are marked with an asterisk (*). 3 Wilcoxon Signed-rank test with Bonferroni correction was applied when values were not normally distributed. 4 Mann–Whitney U-test was applied when values were not normally distributed.
Table 4. Skin hydration values in the PP population, as measured with the Moisturemeter® and expressed as AU. Hydration was measured from the cheek, forearm, and hand every 4 weeks from baseline to Week 12.

<table>
<thead>
<tr>
<th>Location</th>
<th>Time Point</th>
<th>BI-04 (N = 61) Mean (SD)</th>
<th>Placebo (N = 67) Mean (SD)</th>
<th>p-Value 1</th>
<th>p-Value 1</th>
<th>p-Value 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheek, AU</td>
<td>Baseline</td>
<td>44.20 (3.41)</td>
<td>44.33 (3.56)</td>
<td>-</td>
<td>-</td>
<td>0.832</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>44.83 (4.40)</td>
<td>45.74 (5.03)</td>
<td>0.191</td>
<td>0.087</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>43.67 (4.79)</td>
<td>44.82 (4.74)</td>
<td>0.388</td>
<td>1.060</td>
<td>0.213</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>45.00 (3.97)</td>
<td>45.47 (4.41)</td>
<td>0.150</td>
<td>0.049</td>
<td>0.634</td>
</tr>
<tr>
<td>Forearm, AU</td>
<td>Baseline</td>
<td>31.18 (2.98)</td>
<td>31.35 (3.07)</td>
<td>-</td>
<td>-</td>
<td>0.754</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>30.95 (3.07)</td>
<td>31.19 (3.25)</td>
<td>0.470</td>
<td>0.594</td>
<td>0.672</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>30.52 (2.98)</td>
<td>30.61 (3.34)</td>
<td>0.041</td>
<td>0.026</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>30.61 (3.15)</td>
<td>30.53 (3.38)</td>
<td>0.081</td>
<td>0.016</td>
<td>0.600</td>
</tr>
<tr>
<td>Hand, AU</td>
<td>Baseline</td>
<td>38.65 (4.52)</td>
<td>37.96 (4.20)</td>
<td>-</td>
<td>-</td>
<td>0.371</td>
</tr>
<tr>
<td></td>
<td>4 weeks</td>
<td>38.70 (4.53)</td>
<td>38.60 (4.75)</td>
<td>0.903</td>
<td>0.106</td>
<td>0.303</td>
</tr>
<tr>
<td></td>
<td>8 weeks</td>
<td>38.54 (4.20)</td>
<td>37.89 (4.66)</td>
<td>0.754</td>
<td>0.871</td>
<td>0.930</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>38.70 (4.25)</td>
<td>37.74 (4.77)</td>
<td>0.917</td>
<td>0.576</td>
<td>0.644</td>
</tr>
</tbody>
</table>

1 Within-group comparisons to baseline by RM-ANOVA. Significant p-values (p < 0.05) are marked with an asterisk (*). 2 Between-group comparisons at baseline by independent t-test and at Weeks 4, 8, and 12 by RM-ANOVA. Significant p-values (p < 0.05) are marked with an asterisk (*). 3 Wilcoxon Signed-rank test with Bonferroni correction was applied when values were not normally distributed.

Figure 3. Cont.
Figure 3. Skin wrinkle parameters from baseline to 12 weeks in the Bl-04 (black line) and placebo (dashed line) groups in the PP population: average depth of wrinkles, Ra, total wrinkle area, and total wrinkle volume, measured from the corner of the eye. Change from baseline in percentage (CBL %) of (a) average depth of wrinkles, (b) Ra, (c) total wrinkle area, and (d) total wrinkle volume at Weeks 4, 8, and 12 versus the respective baseline and actual values for (e) average depth of wrinkles, (f) Ra, (g) total wrinkle area, and (h) total wrinkle volume in subgroups by age—30–44 years and 45–60 years (international age)—at baseline and Weeks 4, 8, and 12. Figure shows mean ± standard error of the mean (SEM); significant differences between Bl-04 and placebo (p < 0.05) are marked with an asterisk (*).

Statistically significant differences between the Bl-04 and placebo groups were observed in wrinkle parameters that were related to general roughness in the corner of the eye at Week 4 (Table 2, Figure 3).

The average depth of wrinkles increased to a lesser extent from baseline (p = 0.030) with Bl-04 (+2.30%) versus placebo (+6.66%) at Week 4 (Figure 3a). At 8 weeks within the Bl-04 group, the average depth was increased further from baseline, but began to decline by Week 12 (Table 2, Figure 3a). Within placebo, the increase from baseline in the average depth of wrinkles had reduced at Week 8 (around to the level of Bl-04), and further at Week 12, not differing from baseline. At Weeks 8 and 12, there were no significant differences between Bl-04 and placebo.

Ra and total wrinkle volume had similar effect patterns as the average depth of wrinkles between Bl-04 and placebo (Table 2). At 4 weeks, the increases from baseline were significantly higher in the placebo group (+6.42%; p = 0.027 and +6.84%; p = 0.023 in Ra and total wrinkle volume, respectively) compared to Bl-04 (+2.46% and +2.17%, respectively) (Figure 3b, d). Placebo effected a significant within-group increase in Ra still at Week 8 versus baseline (p = 0.030) (Table 2).
In contrast to the pattern, Bl-04 significantly decreased ($p = 0.036$) the total wrinkle area by 0.09% from baseline at Week 4, whereas the area rose 0.17% with placebo (Figure 3c). Within Bl-04 and placebo groups, there were no statistically significant differences against baseline at any time point (Table 2).

In the post hoc analysis, in which the study population was subdivided by age into 30–44 years (total subgroup N = 31; Bl-04 N = 18, placebo N = 13) and 45–60 years (total subgroup N = 97; Bl-04 N = 43, placebo N = 54), older participants had more wrinkles and deeper wrinkles than the younger subjects. The benefits of Bl-04 on the average depth of wrinkles (Figure 3e), Ra (Figure 3f), total wrinkle area (Figure 3g), and total wrinkle volume (Figure 3h) were statistically significant at Week 4 compared to placebo in the younger subgroup ($p = 0.017$, $p = 0.016$, $p = 0.021$, and $p = 0.022$, respectively) but not at Week 8 or 12. In the older subgroup, there were no statistically significant differences between Bl-04 and placebo at any time point.

No statistically significant differences were observed in the mean depth of the largest wrinkle, maximum depth of the largest wrinkle, the total length of wrinkles, Ry, or Rz between Bl-04 and placebo (Table 2). Compared to baseline, at Week 4, there were significant within-group increases in Ry in the Bl-04 and placebo groups ($p = 0.010$ and $p = 0.007$, respectively) and in Rz at Week 8 with Bl-04 ($p = 0.019$) and Week 4 with placebo ($p = 0.007$), after which Ry and Rz values reverted to baseline values.

### 3.3.2. Skin Hydration

To determine the effects of Bl-04 on skin hydration, we measured the water content from the cheek, middle of the forearm, and back of the hand using a Corneometer® CM 825 (Table 3) and Moisturemeter® D (Table 4). The corneometer measured hydration at the surface of the skin to a depth of 10 µm of the stratum corneum, whereas the moisturemeter with S15 probe measured up to 1.5 mm of the dermis. With the corneometer, a statistically significant decrease in hydration was observed on the hand in the Bl-04 and placebo groups in the within-group analysis at 4 ($p < 0.001$ and $p = 0.005$, respectively), 8 ($p = 0.002$ and $p = 0.019$), and 12 ($p = 0.001$ and $p < 0.001$) weeks compared to baseline (Table 3). Hydration in the forearm was significantly decreased only at 12 weeks in the placebo group ($p = 0.003$) compared to baseline values. With the moisturemeter (Table 4), hydration decreased on the forearm in the Bl-04 and placebo groups in the within-group analysis at 8 weeks ($p = 0.041$ and $p = 0.026$, respectively) and at 12 weeks in the placebo arm ($p = 0.016$) compared to baseline values. No other significant observations were noted.

### 3.4. Secondary Outcomes

The effects of Bl-04 on TEWL from the cheek, middle of the forearm, and back of the hand were examined with Tewameter® TM300. A significant decrease in TEWL from the forearm was observed in the Bl-04 and placebo groups in the within-group analysis at 4 ($p < 0.001$ and $p = 0.030$, respectively), 8 ($p < 0.001$ and $p = 0.007$), and 12 weeks ($p < 0.001$ and $p < 0.001$) compared to baseline values (Supplementary Material Table S2). Similarly, TEWL from the hand declined at 4 ($p = 0.005$ and $p < 0.001$), 8 ($p < 0.001$ and $p < 0.001$), and 12 weeks ($p < 0.001$ and $p < 0.001$) in these groups versus baseline values. On the cheek, TEWL decreased in the Bl-04 and placebo groups at 8 ($p = 0.047$ and $p = 0.002$, respectively) and 12 weeks ($p < 0.001$ and $p < 0.001$) against baseline values. No other significant observations were noted.

Skin gross elasticity (R2 parameter) at the cheek, examined using Cutometer® MPA580, increased significantly at Week 4 with Bl-04 and placebo ($p < 0.001$ and $p = 0.028$, respectively), Week 8 for placebo only ($p = 0.008$), and Week 12 for Bl-04 only ($p = 0.007$) compared to baseline values (Supplementary Material Table S3). No other significant changes in skin elasticity (R2) were observed.

A statistically significant decrease in skin gloss was observed at Week 4 with placebo ($p = 0.036$) versus baseline value (Supplementary Material Table S4). No other significant differences in skin gloss were noted.
None of the subjective efficacy or usability items differed significantly between the Bl-04 and placebo groups (Supplementary Material Table S5). The proportion of participants who responded positively (4. Somewhat agree/satisfied/easy or 5. Agree/satisfied/easy) on items of subjective efficacy rose in both groups from Week 4 to Weeks 8 and 12 (within-group significance not tested).

### 3.5. Ancillary Analyses

Descriptive statistics were calculated for all recorded weather parameters (Supplementary Material Tables S8 and S9). Nonparametric methods were used to evaluate the differences between and within treatment groups by visit.

Significant changes in average humidity, average temperature, and UV index from baseline were apparent for almost all within treatment comparisons by visit (Supplementary Material Table S10). The 7-day mean, maximum, and minimum temperature and 7-day mode and maximum UV index rose during the study, and average daily humidity was 58% to 59% at baseline and Week 8 and ranged from 52% to 54% at Weeks 4 and Week 12. There were no statistically significant differences in weather parameters between the Bl-04 and placebo groups at any visit.

TEWL from the cheek correlated negatively with 7-day mean temperature in overall (Pearson $r = -0.211, p < 0.0001$), and in the Bl-04 ($r = -0.214, p = 0.0008$) and placebo groups ($r = -0.209, p = 0.0006$) (Figure 4). In the subgroup analysis, this correlation was diluted and just exceeded statistical significance in the 30–44-year-old subgroup (total subgroup N = 31; Bl-04 N = 18, placebo N = 13) ($r = -0.182, p = 0.0430$), whereas in the older subgroup of those aged 45 to 60 years (total subgroup N = 97; Bl-04 N = 43, placebo N = 54), it became more pronounced ($r = -0.220, p < 0.0001$).

![Figure 4. Correlation of TEWL from the cheek and 7-day mean temperature during the 12-week intervention in the Bl-04 and placebo groups. Black and dashed lines represent the correlation of TEWL and temperature in the Bl-04 and placebo groups, respectively. Individual data points are denoted as circles (Bl-04) and crosses (placebo).](image)

This pattern was also seen for the 7-day minimum and maximum temperatures against TEWL ($r = -0.214, p < 0.0001$ and $r = -0.205, p < 0.0001$, respectively); this correlation was more robust in the older age group ($r = -0.185, p = 0.0393$ and $r = -0.165, p = 0.0674$ in 30–44-year-olds and $r = -0.223, p < 0.0001$ and $r = -0.217, p < 0.0001$ in 45–60-year-olds, respectively). No correlation between TEWL and relative humidity was observed (data not shown).
3.6. Adverse Events

Six and five participants in the Bl-04 and placebo groups, respectively, experienced AEs during the study. All AEs were encountered during the consumption of the IP. One AE in the placebo group was classified as a serious adverse event (SAE) and was evaluated as probably not related to the IP. Emergency unblinding was performed for the SAE, resulting in the participant’s exclusion from the blinded analyses. All other AEs were considered mild and definitively unrelated to the IP. Three AEs in the placebo group and four AEs in the Bl-04 group resulted in the discontinuation of the study.

Although several statistically significant changes in the blood safety tests were observed in both groups from the initial screening to the end of the study (Supplementary Material Table S7), none was considered to be clinically relevant.

4. Discussion

Daily intake of Bl-04 for 12 weeks significantly decreased the total wrinkle area at the corner of the eye at 4 weeks compared to placebo, after which there was no significant difference between groups. In addition, with Bl-04, the average depth of wrinkles, Ra, and total wrinkle volume underwent smaller increases at Week 4 versus placebo; at Weeks 8 and 12, there were no significant differences between groups. Further, there was a difference noticed in subgroups by age. Overall, those who received Bl-04 experienced less fluctuation in the measured wrinkle parameters than placebo-treated subjects during the intervention.

The demonstrated immune system modulation benefits of the Bl-04 strain were one of the key characteristics due to which the strain was selected for this study. Bl-04 is rather anti-inflammatory in nature, inducing high IL-10 production and low levels of the pro-inflammatory IL-12 from human peripheral blood mononuclear cells ex vivo and protecting from colitis in vivo [31]; based on these properties, we surmised that this probiotic could act beneficially for the mucosal immune system and the intestinal barrier protection. Greater intestinal permeability has been associated with several health conditions, including autoimmune diseases, liver-related diseases, diabetes, and neurological and gastrointestinal conditions [32,33]. The gut–skin axis comprises the crosstalk between the gut microbiota, gut mucosal barrier, and immune system, which is then translated to a distal site: the skin [6]. Dysbiosis of the gut microbiota have been noticed to associate with several skin diseases, such as acne vulgaris, AD, and psoriasis [34]. Various factors, such as a fatty-acid-rich diet, alcohol, various medications, smoking, and stress, can cause defects in the intestinal barrier and the leakage of luminal components [35], activating the mucosal immune response and inducing the production of pro-inflammatory cytokines, such as tumor necrosis factor α (TNF-α) and interferon γ (IFN-γ), which disrupt the barrier further [36,37]. Conversely, an anti-inflammatory milieu, through its production of growth factors and anti-inflammatory cytokines, such as interleukin 10 (IL-10) and transforming growth factor β (TGF-β), has been suggested to be protective for the intestinal barrier [37].

The exact mechanisms by which gut barrier permeability manifests in the skin are incompletely understood, but the metabolites that are produced by the gut microbiota might function in such processes. Gut microbiota can metabolize nondigestible dietary components and cellular debris and produce beneficial and harmful, even toxic, metabolites as end products, which can affect gut permeability and, ultimately, skin health [34]. Certain metabolites, such as short-chain fatty acids (SCFAs), particularly butyrate, can benefit gut health by serving as an energy source for colonocytes, improving barrier permeability, and having anti-inflammatory properties [38,39]. SCFAs are readily absorbed, and the amount that is absorbed into peripheral circulation is low compared to the colonic lumen and depends on diet [40].

Butyrate has been shown to increase regulatory T cell numbers, which maintain immune system homeostasis in the skin, as in the gut [41,42]. Bifidobacteria do not produce butyrate per se but instead synthesize acetate and lactate, which are converted into butyrate by other colonic bacteria through cross-feeding interactions [43]. Butyrate has also been shown to stimulate in vitro in L929 murine fibroblasts TGF-β1 and collagen production,
which are important in maintaining the dermal extracellular matrix [44]. The short-chain acids that are absorbed into systemic circulation could function peripherally in the skin, as shown for *Bifidobacterium longum* with or without galacto-oligosaccharides in UV-induced wrinkle formation [45], constituting another mechanism through which Bl-04 could function against skin wrinkling and warranting additional studies.

Skin wrinkles are forming as a result of internal and external factors creating changes in epidermal thickness, lack of hyaluronan, flattening of the dermo-epidermal junction, loss and disorganization of collagen due to the upregulation of collagen-degrading matrix metalloproteases (MMPs), degradation of elastic fibers, and decreased water-holding capacity [5,46,47]. In addition, cellular senescence affects skin aging and wrinkle formation through the secretion of pro-inflammatory cytokines, MMPs, and chemokines that can lead to inflammaging—chronic low-grade inflammation that is associated with aging [48].

Parameters describing the features of deeper wrinkles did not show major changes in this study (Table 2). The development of deep wrinkles is a cumulative, long-term process that takes years [49]. In our study, the intervention was 12 weeks; thus, a longer treatment period might be needed for deeper wrinkles.

Few studies have examined the effects of an oral probiotic supplement as a single active ingredient on skin wrinkles and hydration with a similar methodology [50], which makes the comparison of the results challenging. When our results were compared to a study conducted with a similar design and methods [11], we noted that the baseline values of several comparable wrinkle parameters where considerably higher in volunteers in our study than in Lee et al., in our study, 10-point average roughness (Rz) was over 230 µm and a maximum depth of biggest wrinkle was over 150 µm compared to Lee et al. values of Rz under 90 µm and maximum profile valley depth (Rv) less than 80 µm, respectively,—perhaps partially explaining the observations in our study. Younger participants, in whom the beneficial effects of Bl-04 on skin wrinkles over placebo were especially noticed, had fewer and shallower wrinkles than the older subjects (Figure 3e–h). The younger participants were also closer to the population used in the Lee et al. considering the wrinkle status; Ra describing average roughness was in both under 20 µm. Nonetheless, we could observe visual decreases in several wrinkle features also in participants in the older Bl-04-treated subgroup (Supplementary Material Figure S1).

As with the younger subgroup, in subsets of subjects with self-reported deficient facial skin hydration and deficient facial skin sebum (Table 1), wrinkle volume, an average depth of wrinkles, and Ra increased to a lesser extent with Bl-04 than placebo at Week 4 compared to baseline (data not shown). In addition, in a subgroup with more than 1 h of self-reported daily UV exposure, Bl-04 resulted in a decrease in total wrinkle area, and smaller increases total wrinkle volume, and the average depth of wrinkles versus baseline than placebo at Week 4 (data not shown). As in the older age subgroup, no statistically significant differences between Bl-04 and placebo were seen in total wrinkle volume and area, average depth of wrinkles and Ra in subgroups with self-reported normal skin hydration, normal skin sebum, and less than 1 h of daily UV exposure (Table 1) at any time point (data not shown). No statistically significant differences were observed in total wrinkle volume and area, average depth of wrinkles and Ra in skin roughness subgroups—normal or rough—at any time point (data not shown). This may indicate that participants with suboptimal skin conditions might benefit from Bl-04 use.

In this study, we noted increases in many wrinkle parameters in both groups at Week 4, which then declined at varying rates, whereas in Lee et al. [11], these values decreased quite steadily in both groups throughout the intervention. In a study by Fanian et al. [51], which was also performed from winter to spring, reported similar effects as ours: several parameters of skin roughness increased in the middle of the study period in both groups, but the increase was milder in the treatment group. Thus, as Fanian et al. due to the milder increase, we propose that the treatment may protect the skin against dryness during winter. Winter seasons can vary with regard to weather conditions, and few skin health-
related studies have reported weather conditions and the timing of the intervention, further hindering a comparison of the results.

In Korean women, skin scaliness and elasticity correlated with environmental factors, such as air temperature, pressure, and peak precipitation [52]. Short-term exposure to low humidity has been shown to induce aggravation of skin texture and the formation of fine lines [53]. In our study, TEWL did not correlate with 7-day mean relative humidity—overall or when stratifying by age (data not shown). There are no general guidelines on how to evaluate correlations with the weather in skin research. Relative humidity, rather than absolute humidity, is the most common method of measuring moisture content in the air, but the association of skin parameters with absolute humidity should be considered as well [54]. Our study was conducted during the dry winter season, and indoor heating might have further decreased the humidity, affecting additional dryness of the skin and most likely exacerbating fine line wrinkling [55].

Skin barrier function is negatively affected by climate, and there are clear seasonal differences in hydration of the skin, peaking during the summer and bottoming in the winter, causing winter xerosis [56–58]. Skin appearance can be influenced by controlling the hydration of the skin, and dry skin with alkaline pH has been linked to faster persistent wrinkling in an 8-year longitudinal study [49]. The amount of water accumulated in the outermost corneal layer is defined as skin hydration, which depends on natural hygroscopic agents in corneocytes, natural moisturizing factors, and intercellular lipids [1]. Daily insults from the environment, such as low humidity, surfactants, wind, and sun, can lower the water content in the stratum corneum, causing improper desquamation and the appearance of dry, rough, flaky skin [59,60]. Lifestyle factors, including water intake, can also affect the hydration of the skin, especially if an individual has low prior water consumption [61].

Conversely, TEWL is the diffusion of a small proportion of water molecules through the stratum corneum. It can be used as an indirect measure of skin permeability, and because it is affected by factors in the surrounding microclimate, such as environmental humidity, temperature, and airflow, it should be measured under controlled conditions [62–64], as was performed for all skin measurements in our study, including hydration and TEWL. TEWL values have been shown to correlate inversely with hydration in diseased skin, such as in AD and psoriasis [65]. Hydration has been suggested to decrease with aging in Korean females and during the winter season [66,67].

We did not observe significant changes in hydration of the cheek (Table 3), even when participants were stratified by age, which is consistent with Baek et al. [68]. Instead, we noted significant within-group decreases in TEWL from the forearm, hand, and cheek in the BI-04 and placebo groups (Supplementary Material Table S2), as reported by Lee et al., who showed a significant decrease with a probiotic versus placebo [11]. The lack of an effect of the probiotic over placebo in our study might be related to differences in the study populations between ours and the study by Lee et al., as described above. TEWL has been suggested to remain constant or decrease during aging [69], and a decline has also been noted in Korean females with age [66–68]. Environmental temperature correlated negatively with TEWL in the overall cohort and the 45–60-year-old subgroup and to a lesser extent in those aged up to 44 years.

TEWL has been previously shown to be dependent on environmental temperature, becoming higher during winter versus in summer, but there are contradictory results regarding this pattern [70–73]. In addition, humidity, affecting skin dryness at higher temperatures, might have a greater effect on TEWL than temperature itself [71].

It is poorly understood how TEWL affects wrinkle formation. TEWL correlates with skin hydration and skin surface patterns in the short-term [55]; however, because wrinkle formation is a slow process [49], the longer-term effects of increased permeability should be determined to understand this relationship. Skin hydration is a rather understudied topic and warrants further investigation [55], especially with regard to wrinkle formation.

This study was conducted during the global severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2, or COVID-19) pandemic, during which the use of protective masks
was widely recommended for the general public. Recent results demonstrated that only 3-h use of a fabric or disposable medical mask increased skin hydration in the general population, whereas such changes in TEWL and pH were not observed [74]. Even after short-term 1-h or 6-h use, skin temperature, redness, and sebum secretion rose, with varying results for TEWL and no significant change in skin hydration or elasticity in the cheek area [75]. In contrast, long-term use, 6-h daily use for 2 weeks, decreased skin elasticity and increased skin pores and acne, whereas skin temperature, redness, TEWL, and sebum content did not change significantly [76]. Additionally, mask use for 6 h daily for 4 weeks was shown to enhance wrinkle formation and roughness in nasolabial slopes [77].

In our study, skin hydration, TEWL, elasticity, and gloss were measured in the area of the cheek that would be covered by a mask, and it is unknown whether mask use would also have affected wrinkle parameters at the corner of the eye. Despite the pre-measurement stabilization in a temperature and humidity-controlled room, it remains unknown how putative short- or long-term mask use would have impacted the measured values, warranting further studies. In addition, we noted a decrease in skin hydration on the hand in both groups (Table 3), perhaps due to long-term, frequent use of hand sanitizer and hand-washing with soap during the pandemic [78].

We did not observe any statistically significant effects in the data on the subjective questionnaires. Subjective data that are collected with questionnaires are highly prone to placebo effects and, by nature, measure perceived effects; thus, they should not be compared directly with data that are gathered using objective measurements.

This study provides important novel data on the effects of probiotic supplementation on skin health and is the first report to examine the effects of oral consumption of Bl-04 on skin wrinkle and hydration characteristics. The study was a triple-blinded, rigorously conducted, randomized, placebo-controlled trial, with proper inclusion and exclusion criteria. Seasonal weather conditions were collected from the local meteorological office to analyze the study outcomes. Moreover, skin parameters were evaluated with objective measurements using reliable quantitative methods, the use of which remains limited in dermatological studies.

Studies on probiotics encounter the same issues as other interventional studies, in that the participants are non-responders or responders [79]. We suggest that future similar studies would prescreen participants not only for wrinkles but also for either low-grade inflammation or gut permeability, if the mode of action of an IP is proposed to be linked to the gut–skin axis. Low-grade inflammation could be measured, for instance, with high-sensitivity C-reactive protein or interleukins, such as IL-6 [80–82]. Permeability of the gut is ideally measured by the urinary recovery of molecular markers that are administered, such as non-metabolizable sugars, but this approach would require non-eligible participants to consume such substances, increasing the complexity of the study. Noninvasive biomarkers could be used instead, including urinary or serum fatty acid-binding protein, urinary or serum glutathione S-transferase, serum zonulin, and serum lipopolysaccharide or D-lactate that is derived from gut bacteria, but with caution [35,83].

The gut microbiota is unique to each individual, and with microbially derived metabolites, it might also determine which individual is a responder [79]. Bl-04 has been shown to possess gut microbiota modulatory functions, but this activity is debated [24,84]. It remains unknown whether Bl-04 could influence the composition of the gut microbiota, as has been recently suggested for Lactiplantibacillus plantarum HY7714 [85]; however, as a prescreening tool, the use of gut microbiota would be difficult because dysbiosis in the gut microbiota is observed in intestinal and extraintestinal disorders but not in healthy individuals, who presumably have healthy and unique gut microbiota [86]. Biological samples for efficacy parameters were not collected in our exploratory study, but the results warrant collecting and investigating those in future trials to illustrate the interaction between gut and skin.

Our study has several limitations. The trial was conducted at a single center and included only females, which challenges the generalization of our results to males. The study was short—12 weeks—considering the slow development of wrinkles. Further, the
dose of BL-04 of $1.75 \times 10^9$ CFU was over 5-fold lower than in the corresponding Korean study [11], complicating a comparison between these otherwise similar studies. In the stratification by age, the younger group (30 to 44 years) was smaller than the older subgroup (45 to 60 years), raising some statistical uncertainty with the results.

In addition, we did not take into consideration the menopausal status of the participants, as entering into menopause has been shown to accelerate wrinkle formation [49]. This phenomenon could have diluted the potential beneficial effects of BL-04 on skin wrinkles if many of the participants were entering menopause during the study.

Further, based on the instructions that were given to the participants, consumption of other functional foods that contain probiotics was prohibited, but it was not followed by a self-reported diary or by other measures, which would have improved the quality of the study. Participants might have been celebrating the Korean New Year during the intervention, which could have introduced changes to their dietary habits. The study participants were eligible for the study based on the measurement of dry skin on the cheek and subjective visual subjective inspection of skin wrinkles at the corner of the eye. The latter could have been measured with an objective quantitative method to obtain a more homogeneous population.

5. Conclusions

This randomized, placebo-controlled, triple-blinded clinical trial has demonstrated that the consumption of BL-04 ($1.75 \times 10^9$ CFU) for 12 weeks is safe. This probiotic may alleviate facial wrinkling at 4 weeks, especially in the younger population; however, its effects on wrinkle parameters did not differ significantly at later stages of the intervention, when environmental conditions became more favorable for the skin. No statistically significant differences in hydration parameters were detected between BL-04 and placebo during the study. This study suggests that environmental conditions should be considered more carefully in skin intervention trials and that skin condition, age, and unpredictable confounding factors, such as the use of protective masks, may have effects on the measured parameters.

Supplementary Materials: The following supporting information can be downloaded at: [link].

**Funding:** This research was fully sponsored by Danisco Sweeteners Oy (a subsidiary of International Flavors & Fragrances, IFF), Kantvik, Finland. No public funding was received to conduct this trial.

**Institutional Review Board Statement:** This study was conducted in accordance with the Declaration of Helsinki and approved by the Dermapro Ltd. Institutional Review Board (Seoul, Korea; Approval 1-220777-A-N-01-DICN20181, dated 6 August 2020).

**Informed Consent Statement:** Informed consent was obtained from all subjects who were involved in the study.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors wish to thank Tad Stewart, IFF Health and Biosciences, Madison WI, USA, for management of the IP manufacturing; Sujin Ha, IFF Health and Biosciences, Seoul, Korea, for supporting the IP logistics; Jenni Perämäki, 4Pharma, Turku, Finland, for her contribution to the post hoc statistical analyses; and Markus Lehtinen and Arthur Ouwehand, IFF Health and Biosciences, Kantvik, Finland, for their valuable insights to the manuscript content.

**Conflicts of Interest:** This study was funded by IFF, which manufactures and sells Bl-04. The employees of IFF were salaried staff who participated in the design of the study, the analysis and interpretation of the data, writing of the manuscript, and the decision making with regard to publishing the results. Otherwise, the authors declare no conflict of interest.

**References**

18. Forssten, S.; Ouwehand, A.C. Dose-Response Recovery of Probiotic Strains in Simulated Gastro-Intestinal Passage. Microorganisms 2020, 8, 112. [CrossRef]
35. Ferroni, R.; Fiorani, M.; Abdu Rahman, S.; Matteoli, G. Intestinal Permeability, Inflammation and the Role of Nutrients. Nutrients 2020, 12, 1185. [CrossRef]


60. Dabrowska, A.K.; Adhirtar, C.; Siano, F.; Rotaru, G.M.; Derler, S.; Zhai, L.; Spencer, N.D.; Rossi, R.M. In vivo confirmation of hydration-induced changes in human skin thickness, roughness and interaction with the environment. *Biointerphases* 2016, 11, 031015. [CrossRef]


83. Ajamian, M.; Steer, D.; Rosella, G.; Gibson, P.R. Serum zonulin as a marker of intestinal mucosal barrier function: May not be what it seems. *PLoS ONE* 2019, 14, e0210728. [CrossRef] [PubMed]

