Effectiveness of a machine learning-enabled skincare recommendation for mild-to-moderate acne vulgaris: an 8-week evaluator-blinded randomised controlled trial

Misbah Noshela Ghazanfar, Ali Al-Mousawi, Christian Reimer, Benóný Þór Björnsson, Charlotte Boissard, Ivy Lee, Zarqa Ali, Simon Thomsen

Submitted to: Journal of Medical Internet Research
on: May 24, 2024

Table of Contents

Effectiveness of a machine learning-enabled skincare recommendation for mild-to-moderate acne vulgaris: an 8-week evaluator-blinded randomised controlled trial

Misbah Noshela Ghazanfar[1]; Ali Al-Mousawi[2] MD; Christian Reimer[3] MSc; Benóný Þór Björnsson[3] MSc; Charlotte Boissard[3] MSc; Ivy Lee[4] MD; Zarqa Ali[2] MD, PhD; Simon Thomsen[2, 5] MD, PhD

[1]Department of Dermatology, Bispebjerg Hospital Copenhagen DK
[2]NØIE Copenhagen DK
[3]Division of Dermatology, University of California Los Angeles, Los Angeles, Los Angeles US
[4]4Department of Biomedical Sciences, University of Copenhagen Copenhagen DK

Corresponding Author:
Misbah Noshela Ghazanfar

Abstract

Personalised skincare recommendations may be beneficial for treating mild-to-moderate acne vulgaris (AV). This study aimed to evaluate the effectiveness of a novel machine learning approach in predicting the optimal treatment for mild-to-moderate AV based on self-assessment and objective measures. A randomised, evaluator-blinded, parallel-group study was conducted on 100 patients recruited from an online database and randomised in a 1:1 ratio (groups A and B) based on their consent form submission. Groups A and B received customised product recommendations using a Bayesian machine learning model and self-selected treatments, respectively. The patients submitted self-assessed disease scores and photographs after the 8-week treatment. The primary and secondary outcomes were photograph evaluation by two board-certified dermatologists using the Investigator Global Assessment (IGA) scores and quality of life (QoL) measured using the Dermatology Life Quality Index (DLQI), respectively. Overall, 99 patients were screened, and 68 patients (mean age: 27 years) were randomised into groups A and B. IGA scores significantly improved after treatment in group A but not in group B. DLQI significantly improved in group A from 7.75 at baseline to 3.5 after treatment but reduced in group B from 7.53 to 5.3. IGA scores and DLQI significantly correlated in group A, but not in group B. Adverse reactions were reported in group B, but none in group A. Using a machine learning model for personalised skincare recommendations significantly reduced symptoms and improved severity and overall QoL of patients with mild-to-moderate AV, supporting the potential of machine learning-based personalised treatment options in dermatology.

(JMIR Preprints 24/05/2024:60883)
DOI: https://doi.org/10.2196/preprints.60883

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Effectiveness of a machine learning-enabled skincare recommendation for mild-to-moderate acne vulgaris: an 8-week evaluator-blinded randomised controlled trial

Misbah N. Ghazanfar MD, PhD[1], Ali Al-Mousawi MD[1], Christian Riemer MSc[2], Benóný Þór Björnsson MSc[2], Charlotte Boissard MSc[2], Ivy Lee MD[3], Zarqa Ali MD, PhD[1], Simon Francis Thomsen MD, PhD[1,4]

[1]Department of Dermatology, Bispebjerg Hospital, Copenhagen, Denmark

[2]NØIE, Copenhagen, Denmark

[3]Division of Dermatology, University of California Los Angeles, Los Angeles, California, USA

[4]Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark

Correspondence

Misbah Noshela Ghazanfar, MD, PhD

Department of Dermatology, Bispebjerg Hospital

Bispebjerg Bakke 23, DK-2400 Copenhagen NV, Denmark

E-mail: misbah.noshela.ghazanfar.02@regionh.dk

Statements

Statement of ethics

This study involved no medical intervention; therefore, approval from the ethics committee was not required.

Consent statement

Written consent was obtained from all participants before participation in the study.

Conflict of interest

Charlotte Boissard, Benóný Þór Björnsson, and Christian Riemer are employed by NØIE.

With no relation to the present manuscript, Simon Francis Thomsen has received research support from Janssen, LEO Pharma, Novartis, Sanofi, and UCB and has been a speaker/consultant for AbbVie, Almirall, Eli Lilly, Galderma, Janssen, LEO Pharma, Novartis, Pfizer, Sanofi, Symphogen, UCB, and Union Therapeutics. MNG, AA, ZA, and IL have no conflict of interest.

Funding

The study is sponsored by NØIE (See conflict of interest). None of the other authors have any affiliation with NØIE.

Author contributions

ZA and CR conceived of and planned the study. SFT and IL performed dermatological assessments. CR, BB, and CB analysed the results. MNG, AA, CR, BB, CB, ZA, and SFT contributed to the interpretation of results and models. MNG took the lead in writing the manuscript and was supervised by ZA, CR, and SFT. All authors discussed the results and contributed to the final manuscript. All authors have read and approved the final manuscript.

Data availability statement

Data are available upon request.

Acknowledgements: None

Abstract

Personalised skincare recommendations may be beneficial for treating mild-to-moderate acne vulgaris (AV). This study aimed to evaluate the effectiveness of a novel machine learning approach

in predicting the optimal treatment for mild-to-moderate AV based on self-assessment and objective measures. A randomised, evaluator-blinded, parallel-group study was conducted on 100 patients recruited from an online database and randomised in a 1:1 ratio (groups A and B) based on their consent form submission. Groups A and B received customised product recommendations using a Bayesian machine learning model and self-selected treatments, respectively. The patients submitted self-assessed disease scores and photographs after the 8-week treatment. The primary and secondary outcomes were photograph evaluation by two board-certified dermatologists using the Investigator Global Assessment (IGA) scores and quality of life (QoL) measured using the Dermatology Life Quality Index (DLQI), respectively. Overall, 99 patients were screened, and 68 patients (mean age: 27 years) were randomised into groups A and B. IGA scores significantly improved after treatment in group A but not in group B. DLQI significantly improved in group A from 7.75 at baseline to 3.5 after treatment but reduced in group B from 7.53 to 5.3. IGA scores and DLQI significantly correlated in group A, but not in group B. Adverse reactions were reported in group B, but none in group A. Using a machine learning model for personalised skincare recommendations significantly reduced symptoms and improved severity and overall QoL of patients with mild-to-moderate AV, supporting the potential of machine learning-based personalised treatment options in dermatology.

Keywords: Machine learning, personalised skincare, acne vulgaris, dermatology

Introduction

Acne vulgaris (AV) is one of the most common skin disorders, with a peak incidence in adolescence and early adulthood, affecting approximately 85% of individuals aged 12–24 years. Although acne is most prevalent in teenagers, it can emerge at any age. Approximately 25% and 12% of women and men in their 40s, respectively, report experiencing acne, usually accompanied by a high degree of stigma and impairment in quality of life.[1]

Topical treatments, such as retinoids, antibiotics, and combinations of antibiotics and benzoyl peroxide, are usually used for mild-to-moderate AV, however, previous studies have reported a lack of adherence to topical treatments among patients with AV.[2] A wide range of skincare products are available at beauty stores, pharmacies, and online shops. Many of these products are inefficient, and users usually lack knowledge about which ingredients are effective and beneficial for their skin condition.

A Danish skincare company (NØIE) has developed a method for customising skincare products based on in-depth phenotyping and direct feedback loops from over 80,000 patients with a skin condition by collecting clinical data on skin characteristics during an online survey and combining it with dermatological knowledge, feedback from users, and statistical modelling. In 2019, after 3 years of development, the project successfully launched a data model which modelled personalised skincare solutions based on an individual's specific skin and personal needs.

Therefore, this study aimed to evaluate the effectiveness of a novel machine learning approach for predicting the optimal treatment for mild-to-moderate AV based on subjective patient self-assessment and objective measures, including the physician-rated Investigator Global Assessment (IGA) and the patient-rated Dermatology Life Quality Index (DLQI).

Methods

Study design and participants

This evaluator-blinded, randomised controlled parallel-group trial included 100 patients who were randomised into groups A and B on a 1:1 ratio based on their submission of consent forms. It was conducted in accordance with the Consolidated Standards of Reporting Trials statement.[3] Each group was assessed and assigned treatment based on the self-assessed reporting and image/real-life face-to-face interaction: group A used NØIE's Bayesian machine learning model and group B found and chose products themselves (Figure 1). The patients submitted self-assessed disease scores and a standardised set of facial images, known as a collected dataset, as the skin profile after 8 weeks of adherence to treatment. Subsequently, these assessments and images were scored by two independent board-certified dermatologists to evaluate the effectiveness of the two methods.

Recruitment process

Patients were recruited through several online channels affiliated with NØIE and e-mail campaigns targeting NØIE's database. NØIE contacted respondents to determine eligibility via a short survey before the screening process. The eligibility criteria were patients aged 18–40 years with a known diagnosis of AV who were interested in participating in the study from mid-February onwards for an 8-week duration and resided in Denmark, Germany, the Netherlands, Sweden, the United Kingdom, Switzerland, Belgium, or Austria.

All patients were required to submit a high-resolution image of their facial acne, which an employee of NØIE objectively assessed to confirm the accurate diagnosis (AV) and ensure that the disease severity was mild-to-moderate based on lesion counts. Mild AV was categorised as mostly whiteheads and blackheads with a few papules and pustules, whereas moderate acne was categorised as multiple papules and pustules.

The inclusion criteria were healthy female or male patients aged 18–40 years with mild-to-moderate AV and who had not previously used NØIE products. In contrast, the exclusion criteria were pregnancy, breastfeeding, or any changes in birth control during the intervention period since these

would cause fluctuations in the hormonal impact on disease severity. Patients who used prescription medical products/treatments for acne treatment were also excluded.

Ethics approval and informed consent

All patients provided written informed consent before enrolment. Approval from the scientific ethics committee was not required as no medical interventions were conducted in the study.

Bayesian model guidance for product development

NØIE is established upon substantial data collection from over 65,000 individuals with various skin diseases, including skin reactions to well-categorised products commercially sold across Europe. These data were reverse-engineered into a matrix of stratified user segmentation, with an underlying layer of products developed by NØIE to support the optimal needs of individual user's skin. Additionally, this precision medicine approach has only been used in oncology to date; however, NØIE has incorporated Bayesian modelling to stratify users based on differences in their epigenetic features, lifestyle, and personal preferences and their response to the products being used, based on the collected data, in-house product development, and direct contact with each user. The feedback from real-world data not only trains the model for precision but also guides skincare product development simultaneously to better meet the diverse needs of patients.

Precision medicine is an emerging approach in clinical research and patient care which focuses on understanding and treating diseases by integrating multimodal or multiomics data from each patient to make personalised treatment decisions.[4] Additionally, dealing with the large and intricate datasets generated by precision medicine diagnostics requires the development of innovative techniques to process and interpret this complex information. Concurrently, rapid advancements in computer science have enabled the storage, processing, and analysis of these intricate datasets—a task that traditional statistics and early computing technologies could not accomplish. Therefore, the Bayesian

modelling approach provides a means to formalise prior beliefs and combine them with available observations, aiming to derive rational criteria for optimal decision-making and measure the outcomes of these decisions.[4]

This approach forms the foundational core of NØIE, aiming to identify intricate patterns in data for making predictions and classifications and conducting advanced exploratory data analysis on new, unseen data to guide their product development and distribution for better and safer treatments for various phenotypes of a given skin disease.

The fundamental input for the Bayesian model was the patient undergoing skin test, which is a cumbersome in-depth online survey capturing 31 parameters relevant to the underlying skin condition or disease. These different data points are stored in real-time, forming the patient's skin profile, which the Bayesian models activate to recommend the ideal skin products for providing alleviation. Since all patients are asked to provide feedback on the effectiveness of their given treatment, it establishes a closed loop in the modelling process where results are automatically considered, thereby continuously strengthening the model's precision.

Product description

Since NØIE's products are classified as cosmetics and medical device class I, the ingredients distinguish themselves from classic active pharmaceutical ingredients known in the pharmaceutical industry, although they rely on synergistic effects between conventionally used cosmetic ingredients and innovative modified peptides. Conventional ingredients include salicylic acid, retinol, and niacinamide, whereas innovative ingredients include *Curcuma longa* callus lysate, *Morinda citrifolia* callus culture, and Lactiplantibacillus fermentation lysate. Both groups' routine skincare regimens comprised a face cleanser and cream.

Group A: Customised skincare

All patients in this group underwent a skin test to create a skin profile. Using the Bayesian model, they received product recommendations including a personalised face cleanser and cream. These products were shipped by NØIE.

Group B: Self-selected skincare

Similar to group A, patients in group B filled out using the same skin profile as their starting point. However, compared with group A, it did not activate NØIE's machine learning endpoint for a recommendation. All patients in this group were instructed to select a face cream and cleanser that they believed would improve their acne symptoms over 8 weeks. The choice of product was completely at the patients' discretion; they were permitted to seek advice from family, friends, pharmacists, and doctors but were not allowed medical treatment. NØIE purchased and shipped the self-selected products to patients, or patients purchased the products themselves and NØIE refunded the invoice.

Adverse reactions

Patients were instructed to contact NØIE immediately in the event of any adverse reactions. In such cases, patients were given the option to either substitute the product or were automatically assigned an alternative formulation by NØIE's underlying machine learning model.

Dermatological assessment

A dermatological assessment was conducted based on three images provided by the patients on days 0 and 56. The three images included each cheek side and a frontal profile captured in high resolution, with balanced lighting and without makeup (Figure 2).

The two assessors were licenced dermatologists located in Denmark and the United States of

America, respectively, and unaffiliated with NØIE in any sense. Additionally, the images were rated according to the IGA,[5] and the assessors were blinded to each other's ratings and the group origin of the assessed participant alongside the three images on days 0 and 56. Therefore, biasing the outcomes of the assessments was impossible.

Primary and secondary outcomes

The primary outcome measure was the evaluation of changes in acne severity based on IGA scores by blinded assessors, whereas the secondary outcome was quality of life measured using the DLQI.IGA scores were graded from 0 to 4 [from clear (0) to severe acne (4)], and only facial acne was included in the assessment. Chest, back, and shoulder acne was not considered in this study. Both inflammatory and non-inflammatory lesion counts have been reported.[5] Furthermore, the DLQI is a 10-item retrospective questionnaire which assesses the quality of life of patients with skin conditions. Each question has scores ranging from 0 to 3, with a maximum total score of 30. A higher score indicates a greater impairment in quality of life, and the minimal clinically important difference for DLQI is 3.10 points.[6]

Self-assessment

The self-assessment conducted during the study covered five lesion types for the patients. Additionally, the question '*Do you deal with symptom X*' comprises a response scale ranging from '*Not at all*' to '*Extremely*'. Subsequently, the positioning on the sliding scale was automatically converted to a numerical grade in NØIE's backend, where direct '*Not at all*' and '*Extremely*' are equivalent to 0 and 10, respectively. This implies that patients could score their acne severity from 0 to 50 across all five lesion types. The five acne lesion types assessed were papules, pimples, blackheads, whiteheads, and oiliness, excluding nodules or cysts as they mainly belong to the severe state of AV.

Changes in hypothetical medical intervention from the blinded dermatologists

A third-party unaffiliated dermatologist with decades of experience in treating AV across all severities was randomly assigned to the group receiving customised treatment (group A) according to the sequence and the order of the intervention. Based on the three photographs, we assessed which treatment to prescribe, if any, as a hypothetical action which could be considered a development caused by the study intervention, as all other parameters remained unchanged. Supplementary Table 1 presents the six interventions performed by a blinded dermatologist.

Intervention severity was scored from 1 (least severe) to 6 (most severe). The mean severity was calculated at baseline and after the 8-week intervention, as well as the mean difference for each patient. The higher the mean score, the more severe the intervention required. The difference was calculated as the intervention score after the 8-week treatment minus the intervention score at baseline, where a negative score reflects a downgrade in the severity of the intervention required.

A t-test was performed to determine the difference between the interventions at baseline and after 8 weeks. Statistical significance was set at an alpha value of 0.05.

Statistical analysis

Differences in mean values within and between groups were analysed using the t-test and by comparing percentage differences. Correlations were calculated using Pearson's correlation coefficients. Statistical analyses, data handling, Pearson correlation calculations, and graphing were performed using Python version 3.7.12, Pandas library version 1.3.5, Numpy library version 1.21.6, and Seaborn library version 0.12.1, respectively.

Cohen's test to ensure adequate sample sizes

A power calculation using Cohen's test was conducted before initiating the recruitment process to estimate the sample size required for each group for determining the likelihood of detecting an effect

in the experiment if it truly existed. First, the effect size (ES) was calculated by substituting the proportions of patients expected to be improved by each treatment, $p_1 = 0.86$* and $p_2 = 0.50$**, and the overall proportion, p = 0.68 [i.e. (0.86 + 0.50) / 2]:

$$ES = \frac{|p_1 - p_2|}{\sqrt{p(1-p)}} = \frac{|0.86 - 0.5|}{\sqrt{0.68(1-0.68)}} = 0.7717$$

*$p_1 = 0.86$ refers to the product satisfaction rate obtained by NØIE within the last 2 years. **$p_2 = 0.5$ was set as a high estimate for the self-selected group since competent guidance was expected.

Subsequently, we calculated the sample size (n_i) for each group (i = 1, 2). ES was used in the equation for two independent samples, along with the confidence level, ensuring that a two-sided test with a 5% level of significance (i.e. α = 0.05) and 80% power to detect the estimated response difference between the two groups. The equation is presented as follows:

$$n_i = 2\left(\frac{Z_1 - \frac{\alpha}{2} + Z_1 - \beta}{ES}\right)^2 = 2\left(\frac{1.96 + 0.84}{0.7717}\right)^2 = 26.32$$

$n_1 = 26$ and $n_2 = 26$ ensure that the test of the hypothesis has 80% power to detect a product effect. We anticipated patient dropouts due to strict inclusion criteria and non-compliance with product usage. Therefore, we recruited 50 patients in each group ($n_1 = 50$ and $n_2 = 50$).

Results

Randomisation and screening

A total of 99 patients were screened for randomisation. The failure to obtain the intended 100 patients was attributed to a strict deadline, given the coordination required for product shipment and the need to minimise the period between the initial screening of AV severity and the first day of intervention.

Of the 99 patients, 19 were excluded before randomisation because of the inclusion criteria. Six patients in each group were excluded from data analysis due to non-compliance with usage guidelines (self-reported) or intervention, as shown in Figure 1.

The patients excluded after filling out the initial dataset did not exhibit any specific characteristics regarding age, sex, or disease severity compared with those analysed in either of the two groups. Although modafinil and lamotrigine are not prescribed for AV, they are well-known for their ability to affect the skin and were among the medications causing exclusion due to their potential interference with the study outcome.

In total, 68 patients were randomised into group A (customised), which included 36 patients (34 females and 2 males) with an average age of 27.1 years, and group B (self-selected) comprising 32 patients (30 females and 2 males) with an average age of 26.3 years (Table 1). The duration of AV was 5 and 6 years in groups A and B, respectively. Additionally, 58.3% and 62.5% of the patients in groups A and B, respectively, had previously undergone medical treatment. Baseline DLQI were 7.75 and 7.53 in groups A and B, respectively, indicating a moderate effect on quality of life in both groups.

Primary outcome: IGA assessments

The mean IGA score improved from 1.53 ± 0.83 [±standard deviation (SD)] at baseline to 1.21 ± 0.76 after the 8-week treatment in group A (mean difference in IGA score = 0.32, 95% confidence interval [CI] = 0.05–0.58, p = 0.02) and improved from 1.42 ± 0.42 at baseline to 1.33 ± 0.87 after the 8-week treatment in group B (mean difference in IGA score = 0.09, 95% CI = -0.19 to 0.38, p = 0.51) (Table 2).

Table 3 illustrates the agreement between the two assessors' IGA scores. The two assessors disagreed in 36 cases but agreed upon no change (16), improvement (13), and regression (3) in 32 cases.

Significant correlations were observed between the two assessors at baseline (r = 0.44, 95% CI = 0.23–0.62, p < 0.001) and after the 8-week treatment (r = 0.54, 95% CI = 0.35–0.69, p < 0.001).

Secondary outcome: quality of life

This parameter measured using the DLQI significantly improved in group A from 7.75 ± 5.03 at baseline to 3.5 ± 4.1 after the 8-week treatment (mean difference in DLQI = 4.3, 95% CI = 2.1–6.4, p < 0.001). In contrast, a reduction in DLQI from 7.53 ± 6.16 at baseline to 5.3 ± 4.7 after the 8-week treatment was observed in group B (mean difference in DLQI = 2.3, 95% CI = -0.50 to 5.0, although this was not significant (p > 0.05) (Table 2).

Correlation between the IGA score and DLQI

Weak and moderate correlations were observed between the IGA score and DLQI at baseline (r = 0.09, p > 0.05) and after the 8-week treatment (r = 0.532, p < 0.001) in group A, respectively. However, a weak correlation was observed between acne severity and DLQI at baseline (r = 0.173, p > 0.05) and after the 8-week treatment (r = 0.146, p > 0.05) in group B (Figure 3).

Self-assessed symptoms

Group A had a total mean symptom score of 26.6 ± 4.9 at baseline which improved to 20.3 ± 7.6 after the 8-week treatment (mean difference = 6.3, 95% CI = 3.22–9.32, p < 0.001). In contrast, group B had a total mean symptom score of 24.9 ± 7.6 at baseline, which improved to 20.6 ± 8.44 after the 8-week treatment (mean difference = 4.31, 95% CI = 0.26–8.37, p = 0.03) (Table 2). Furthermore, significant reductions in the proportions of blackheads (25%), whiteheads (31%), pimples (26%), and skin oiliness (23%) were observed in group A after the 8-week follow-up. In total, 61% and 56% of the patients in groups A and B, respectively, reported that their skin appeared healthier after using the customised skincare routine and the 8-week treatment, respectively. Greater

reductions in the proportions of blackheads (53% vs 35%) and pimples (67% vs 34%) were observed in group A than in group B. Moreover, A significant visual improvement in the skin was observed in group A compared with that in group B (67% vs 37%, p = 0.008).

Correlation between self-assessed symptoms and DLQI

Weak and moderate correlations were observed between acne severity (self-assessed) and DLQI at baseline (r = 0.264, p > 0.05) and after the 8-week treatment (r = 0.469, 95% CI = 0.16–0.68, p = 0.004) in group A, respectively. In group B, a weak correlation was observed between acne severity and DLQI at baseline (r = 0.203, p > 0.05) and at the 8-week follow-up (r = 0.280, p > 0.05) (Figure 4).

Adverse reactions

Overall, three patients in group B reported minor adverse events in the form of blushing, stinging, and itchy sensations upon the first application on day 1. They contacted NØIE and were instructed to discontinue the products and find alternative products the same day. Notably, the alternative products were well tolerated. However, none of the patients in group A reported any adverse events.

Changes in hypothetical medical intervention from the blinded dermatologists

The mean severity score in group A at baseline was 2.97 ± 1.77, which improved to 2.22 ± 1.66 after 8 weeks of intervention (mean difference=0.75 ± 1.84; (95% CI = 0.23–1.27, p = 0.02). After the 8-week intervention, 42.9% of the patients were administered only dermo-cosmetics. Supplementary Figure 1 illustrates the need for a hypothetical medically prescribed treatment for AV at baseline and after the 8-week follow-up for each intervention.

Discussion

Precision medicine is an emerging approach in the field of dermatology, and recent clinical research has shown that the impact and management of skin diseases differ among patients. This has resulted in an increased focus on the development of personalised treatment approaches to optimise treatment response, minimise adverse reactions, and improve overall quality of life of patients.[7,8] However, the knowledge of optimal personalised treatment approaches in clinical practice remains limited.

Main findings

This 8-week, evaluator-blinded, randomised controlled trial, based on patient-taken photographs, evaluated the effectiveness of a novel machine learning approach for predicting the optimal treatment for mild-to-moderate AV, using dermatological evaluation combined with self-assessment. We found that IGA scores, as assessed by board-certified dermatologists, demonstrated a significant reduction in acne severity in group A compared with that in group B, indicating that personalised product recommendations generated by a machine learning model were more effective in improving acne severity than a self-selected treatment approach.

Furthermore, a significant improvement in acne severity (self-assessed) was observed after the 8-week treatment in both groups. Significant reductions in acne symptoms, such as blackheads, whiteheads, pimples, and skin oiliness, were observed in group A compared with that in group B. Therefore, this finding adds to the potential benefits of personalised treatment approaches for patients with AV.

Interpretation

Previous studies have reported that even mild acne significantly impacts the psychological well-being and quality of life of patients.[9] In our study, a significant improvement was observed in the quality of life of patients in group A after the 8-week treatment but not in group B. This suggests that the improvement in quality of life was greater with personalised product recommendations than with

self-selected products. Furthermore, a weak correlation was found between acne severity and DLQI in both groups, suggesting that the impact of acne on quality of life is weakly correlated with its severity, which highlights the complexity of clinical assessments and the psychological impact of acne. These findings align with those of previous reports showing no correlation between acne severity and quality of life impairment.[10,11]

Minor adverse reactions, including blushing, stinging, and itchiness, alone were reported from the use of self-selected products. However, no minor or severe adverse reactions were reported from the use of customised treatment. Therefore, this finding indicates that personalised treatments may have a better safety profile than self-selected products.

Strengths and limitations

The strengths of this study include its randomised design which ensured the assigning of unbiased treatment and the relatively large group of patients. Furthermore, two board-certified dermatologists assessed disease severity using the IGA score. A substantial agreement was noted between the IGA assessments of the two board-certified dermatologists regarding the dynamics of acne evolution throughout the trial. However, IGA scores are limited to the face and do not cover disease activity on the back or chest. Changes in the IGA score/value are small, making it difficult to be used in clinical settings with interventions.[5] The Global Acne Grading System, another score which might be more suitable for research purposes, is a broader scoring system and provides a more detailed picture of disease severity; however, it is time-consuming and difficult to use in clinics with limited time.[12] This study limitations include a short follow-up time (8 weeks), primary focus on mild-to-moderate AV, and self-assessment of acne severity. Furthermore, self-assessment methods have been reported to be unreliable because it is difficult for patients to objectively score the severity of their acne.[13] Further research is needed in personalised skincare and machine learning models with longer follow-up time.

Conclusion

The use of a machine learning model for personalised skincare recommendations significantly improved the severity of acne, reduced symptoms, and improved the overall quality of life of patients with mild-to-moderate AV. Therefore, these findings support the potential of machine learning-based personalised treatment options in dermatology.

References

1. Oge' LK, Broussard A, Marshall MD. Acne vulgaris: diagnosis and treatment. Am Fam Physician 2019;100:475-484.
2. Sevimli Dikicier B. Topical treatment of acne vulgaris: efficiency, side effects, and adherence rate. J Int Med Res 2019;47:2987-2992. https://doi.org/10.1177/0300060519847367
3. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;24:18. https://doi.org/10.1016/j.ijsu.2010.09.006
4. Balaneshinkordan S, Kotov A. Bayesian approach to incorporating different types of biomedical knowledge bases into information retrieval systems for clinical decision support in precision medicine. J Biomed Inform 2019;98:103238. https://doi.org/10.1016/j.jbi.2019.103238
5. Alsulaimani H, Kokandi A, Khawandanh S, Hamad R. Severity of acne vulgaris: comparison of two assessment methods. Clin Cosmet Investig Dermatol 2020;13:711-716. https://doi.org/10.2147/CCID.S266320
6. Basra MKA, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. Br J Dermatol 2008;159:997-1035. https://doi.org/10.1111/j.1365-2133.2008.08832.x
7. Šitum M, Franceschi D, Franceschi N. Challenges and strategies in dermatologic therapy—Personalized medicine, patient safety, and pharmacoeconomics. Dermatol Ther 2019;32:e13011. https://doi.org/10.1111/dth.13011
8. Nie D, Tegtmeyer K, Zhao J, Lio PA. Developing patient-specific adverse effect profiles: the next frontier for precision medicine in dermatology. J Dermatolog Treat 2020;31:211-212. https://doi.org/10.1080/09546634.2019.1687810
9. Gieler U, Gieler T, Kupfer JP. Acne and quality of life - Impact and management. J Eur Acad

Dermatology Venereol 2015;29:12-14. https://doi.org/10.1111/jdv.13191

10. Gupta A, Sharma YK, Dash KN, Chaudhari ND, Jethani S. Quality of life in acne vulgaris: relationship to clinical severity and demographic data. Indian J Dermatol Venereol Leprol 2016;82:292-297. https://doi.org/10.4103/0378-6323.173593
11. Ilgen E, Derya A. There is no correlation between acne severity and AQOLS/DLQI scores. J Dermatol 2005;32:705-710. https://doi.org/10.1111/j.1346-8138.2005.tb00829.x
12. Hacivelioglu S, Gungor ANC, Gencer M, Uysal A, Hizli D, Koc E, Cosar E. Acne severity and the Global Acne Grading System in polycystic ovary syndrome. Int J Gynecol Obstet 2013;123:33-36. https://doi.org/10.1016/j.ijgo.2013.05.005
13. Adah R, Yusufu H, Otene QAV. Epidemiology and perception of acne among adolescents in Jos, Nigeria: cross-sectional school-based study. JMIR Dermatol 2023;6:e44441. https://doi.org/10.2196/44441

Table 1: Patient characteristics.

	Group A (Customised)	Group B (Self-selected)
Age, mean, years (range)	27.1 (20–38)	26.3 (19–35)
Sex, n (%)		
26- Female	34 (94.4)	30 (93.7)
27- Male	2 (5.5)	2 (6.3)
Duration of AV, mean (SD) years	4.94 (3.75)	5.98 (4.84)
Previous use of acne medication, n (%)	58.3	62.5

Abbreviations: AV, acne vulgaris; SD, standard deviation

Table 2 : Primary and secondary outcomes at baseline and 8-week follow-up.

	Baseline	**Follow-up at 8 week**	**p-value**
IGA, mean ± SD			
- group A	1.53 ± 0.83	1.21 ± 0.76	0.02
- group B	1.42 ± 0.42	1.33 ± 0.87	0.51
DLQI, mean ± SD			
- group A	7.75 ± 5.03	3.5 ± 4.1	< 0.001
- group B	7.53 ± 6.16	5.3 ± 4.7	> 0.05
Self-assessment, mean ± SD			
- group A	26.6 ± 4.9	20.3 ± 7.6	< 0.001
- group B	24.9 ± 7.6	20.6 ± 8.44	0.03

Abbreviations: IGA, Investigator Global Assessment; DLQI, Dermatology Life Quality Index; SD, standard deviation

Table 3: Matrix showing inter-agreements plotted as aggregate numbers for the two groups.

Assessor evaluation	Assessor #2 IGA score reduced	Assessor #2 IGA score remained unchanged	Assessor #2 IGA score increased
Assessor #1 IGA score reduced	13	5	3
Assessor #1 IGA score remained unchanged	9	16	7
Assessor #1 IGA score increased	6	6	3

Colour coding: green, consensus between two assessors; yellow, minor disagreement; red, contradictory assessments. Abbreviation: IGA, Investigator Global Assessment

Figure legends

Figure 1: The Consolidated Standards of Reporting Trials diagram.

Abbreviation: IGA, Investigator Global Assessment

Figure 2: Patient photographs for assessment by dermatologists.

Figure 3 Correlation between acne severity (IGA score) and DLQI before and after intervention in groups A and B.

Abbreviations: DLQI, Dermatology Life Quality Index; IGA, Investigator Global Assessment

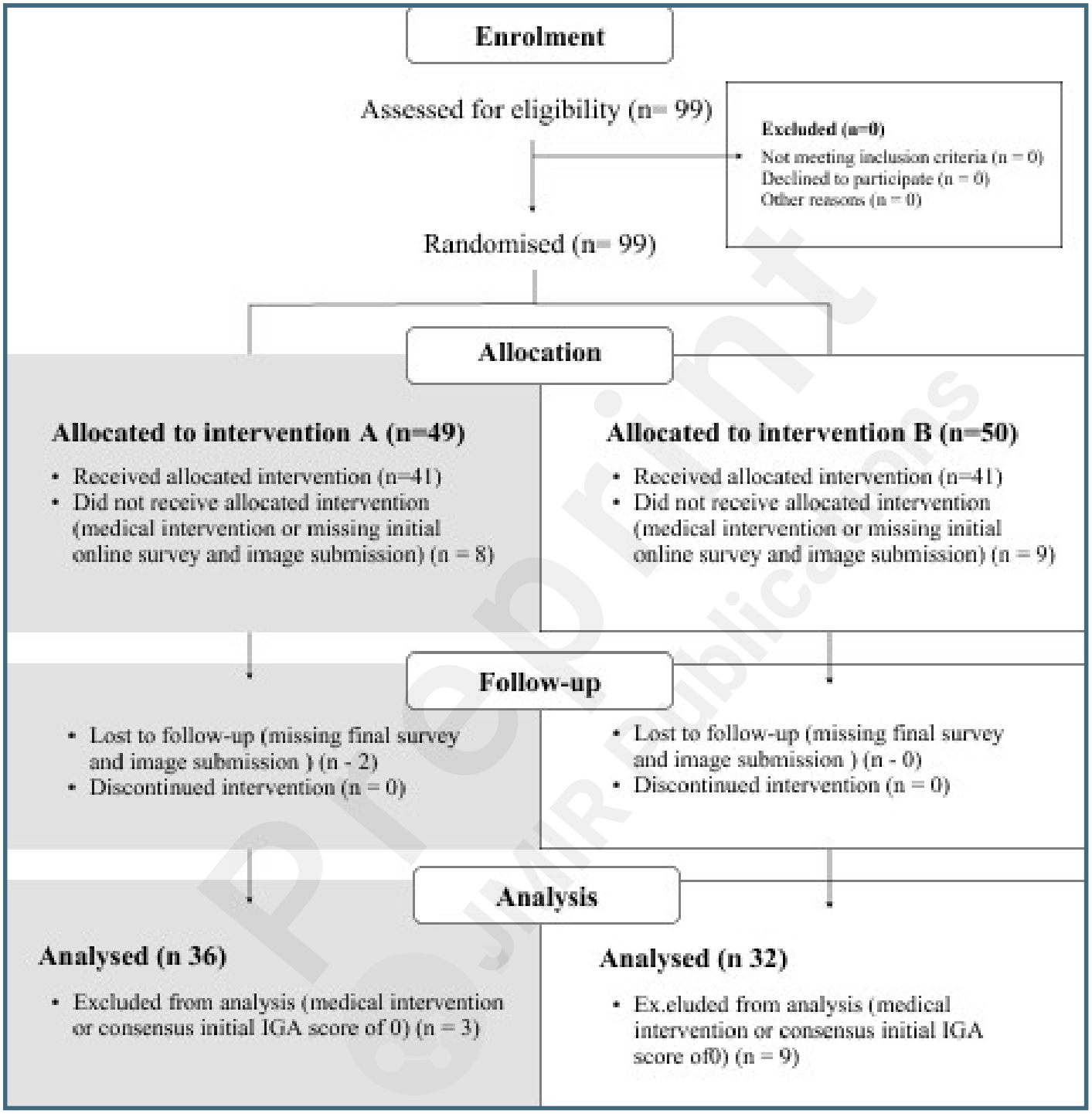
Figure 4: Correlation between acne severity (self-assessed) and DLQI before and after intervention in groups A and B.

Abbreviations DLQI, Dermatology Life Quality Index

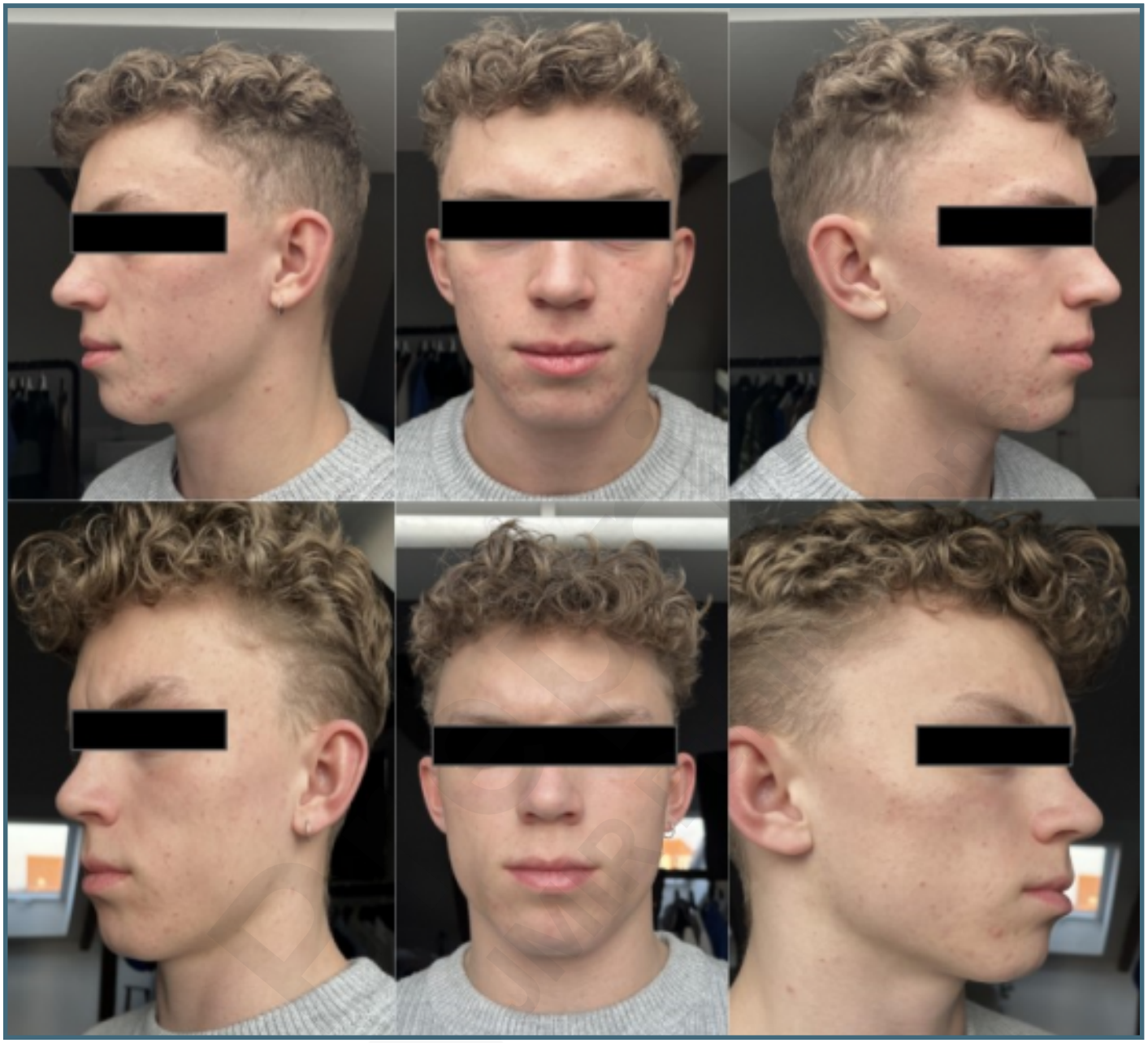
Supplementary Files

Figures

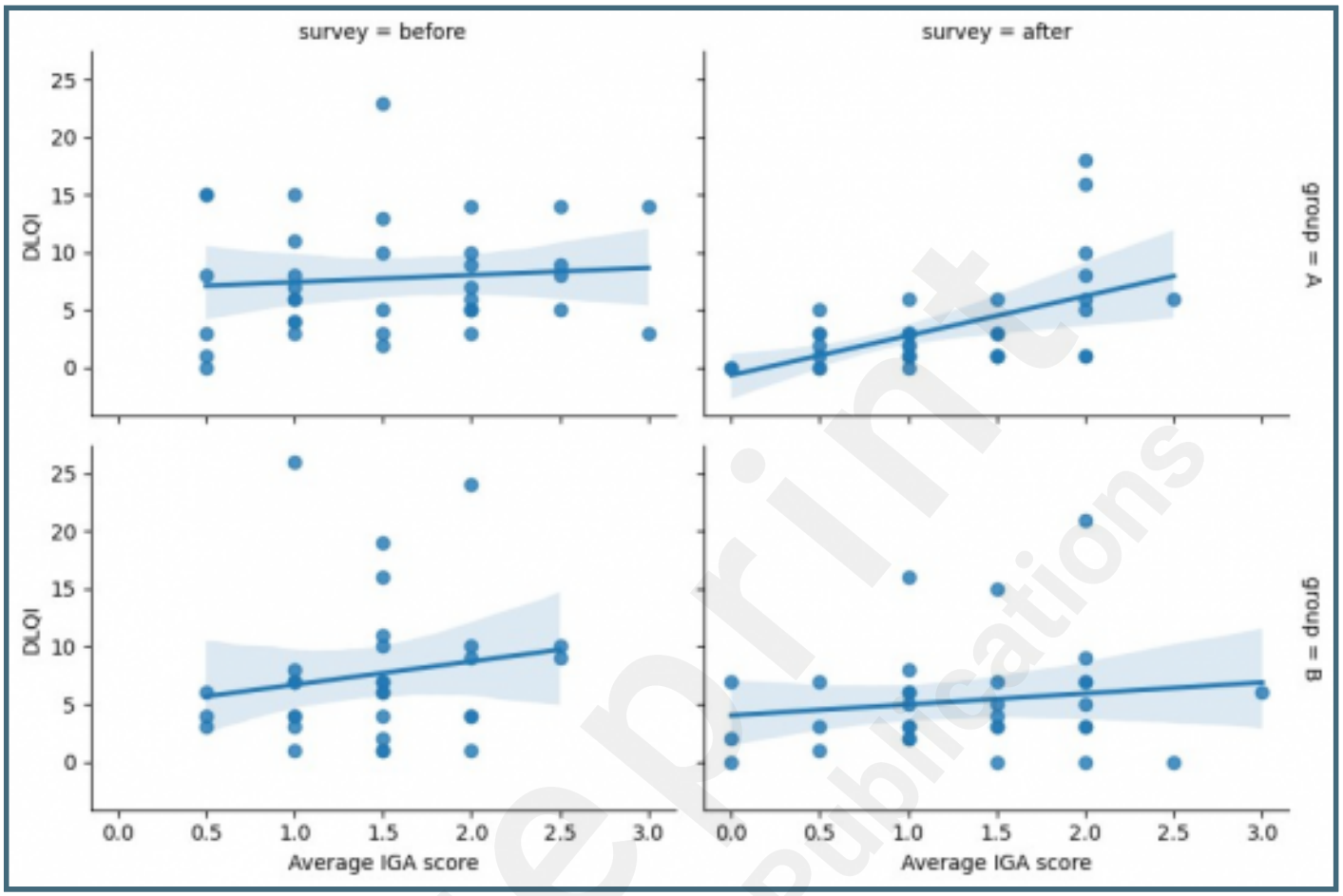
CONSORT diagram.



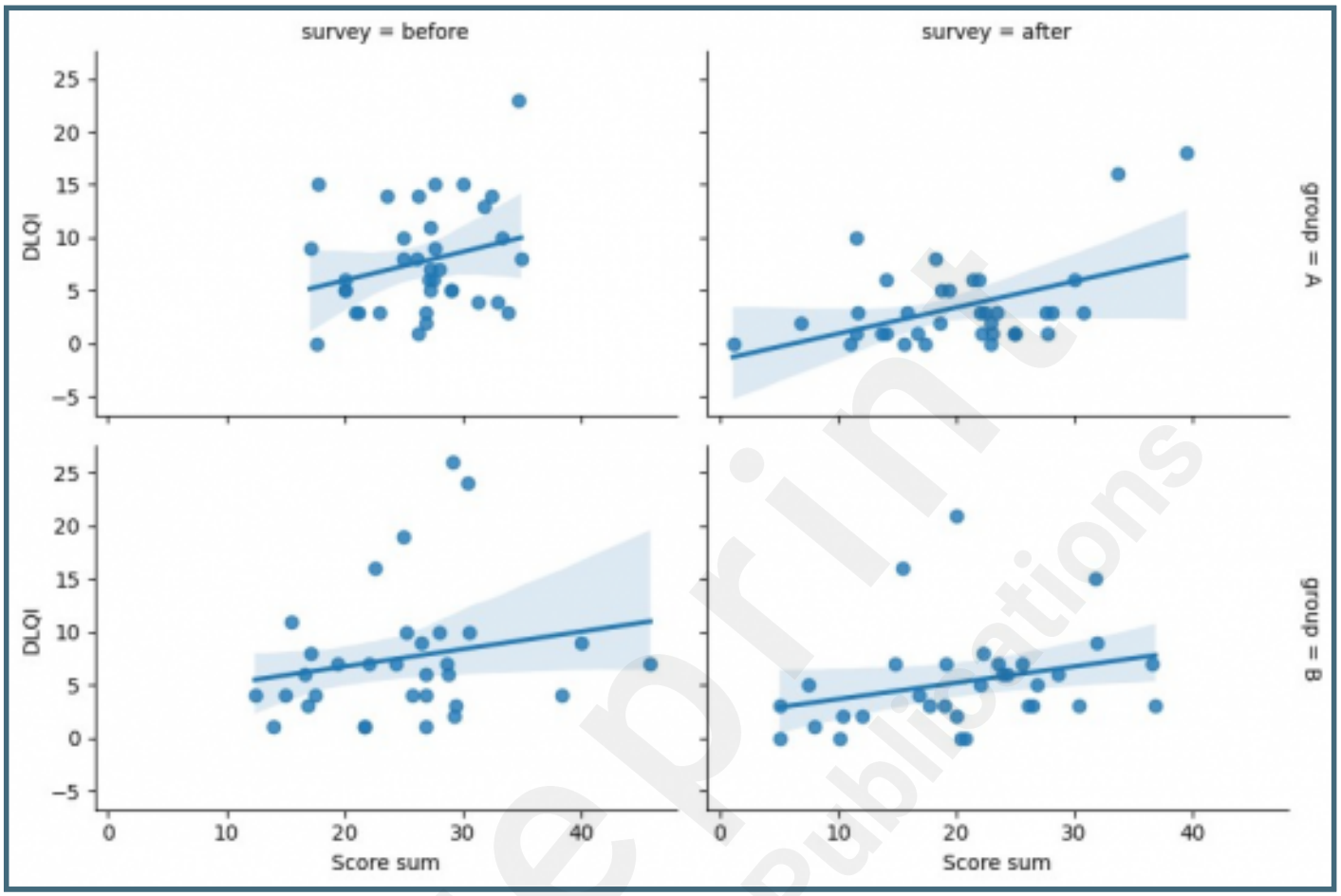
Photos taken of a patient for assessment by dermatologists.



Correlation between acne severity (IGA) scores and DLQI before and after intervention in group A and B.



Correlation between acne severity (self-assessed) and DLQI before and after intervention in group A and B.



CONSORT (or other) checklists

Consort.
URL: http://asset.jmir.pub/assets/48ca229cc50713aa0125e3e8310d61d5.pdf