International Journal of Research in Dermatology
Merja AM et al. Int J Res Dermatol. 2023 Sep;9(5):262-268
http://www.ijord.com

DOI: https://dx.doi.org/10.18203/issn.2455-4529.IntJResDermatol20232540

Original Research Article

Establishing reliability of dermatologist trained and validated evaluators for assessments of dermatological parameters: a comprehensive overview of concepts and techniques

Apeksha M. Merja, Nayan K. Patel, Maheshvari N. Patel*, Megha K. Yadav, Nistha A. Jani

NovoBliss Research Pvt. Ltd., Ahmedabad-Gandhinagar, Gujarat, India

Received: 03 July 2023
Revised: 31 July 2023
Accepted: 17 August 2023

***Correspondence:**
Maheshvari N. Patel,
E-mail: maheshvari@novobliss.in

ABSTRACT

Background: Multi-centered, multi-visit outpatient dermatology studies provide participant-friendly convenience, drawing diverse volunteers while avoiding extended facility stays. However, maintaining a consistent dermatologist for such studies poses challenges, possibly leading to incomplete clinical readings and delaying accurate data collection from predefined datasets.
Methods: Led by a Dermatologist, the training aimed to ensure harmony in assessments among trained evaluators when utilizing digital photos to analyze dermatological parameters. It aimed to establish accord between the Dermatologist and Trained Evaluators for inter- and intra-observer evaluations. The training encompassed multiple parameters, such as Draize Scale, Fitzpatrick Skin Type, Glogau Skin Age Classification, IGA (Acne), PGA with Griffith Scale, and PGA for brittle nail signs.
Results: The results of the statistical analysis for photographic evaluation inter-evaluator reliability were evaluated. Based on the data, a strong correlation was observed between the Dermatologist's dataset and the Trained Evaluator's dataset. The Pearson chi-square test yielded a p-value of <0.05, indicating significant correlation among the trained evaluators' data compared to the dermatologist's data. All evaluators met the 80% acceptance criteria outlined in the training plan. Both evaluators and dermatologists were deemed suitable for dermatological assessments and photographic evaluations in relevant studies.
Conclusions: Theoretical training enabled evaluators to comprehend various dermatological parameters, including Draize Scale, Fitzpatrick Skin Type, Glogau Skin Age Classification, PGA with Griffith Scale, IGA (Acne), and PGA for brittle nail signs. Trained evaluators are now qualified to independently serve as "Dermatologist Validated and Trained Evaluators" in future studies. This approach is applicable for multi-centered, multi-visit dermatological clinical studies.

Keywords: Reliability, Acne, Brittle nails, IGA, Validation, Glogau skin age

INTRODUCTION

Dermatology is a branch of medicine that deals with the skin, hair and diseases of the skin and hair. It involves the study, research, diagnosis, and management of any health conditions that may affect the skin, fat hair, nails, and membranes.[1] Dermatology is primarily a visual discipline. Dermatology competences should include, in addition to effective clinical interviewing and detailed descriptions of skin lesions, appropriate management (diagnosis,

differentiation, and treatment/therapy) of common skin disorders.

In past, when there were no bio instruments available to evaluate the dermatological condition, dermatologists, used to evaluate the condition visually with scoring scales.[2] In the present scenario, there are a number of techniques and bio instruments available. When applicable, bio-instrumental measurements are also useful because of their objectivity and quantification of the results. However, visual and clinical evaluation, if performed by an expert, allows good qualitative and quantitative evaluation.[3] Both visual and bio-instrumental evaluations are crucial to a comprehensive evaluation.

As per the present scenario, one of the major challenge is to manage the availability of the same dermatologist for the controlled multicentric, multi-visits/timepoints, clinical trials requires longer treatment duration with several clinical visits and frequent clinical readings (i.e. hourly reading). Due to the unavailability of the same dermatologist during such controlled clinical trials, many times the clinical readings get missed resulting into loss of getting accurate data in a timely manner with predefined datasets.

Hence, we as NovoBliss Contract Research Organization decided to develop a novel aspect to avoid such a scenario in view of having accurate, timely readings from the trial patients/subjects of cosmetics, and personal healthcare products, by the "Dermatologist's Trained and Validated Evaluators" and those to be doing subjective scoring referring to FDA guidelines allowing to have trained evaluators for skin irritation, sensitization scoring, skin blanching, multi-centric clinical studies with many other skin attributes.

Objectives

The primary goal of this training was to determine the consistency between different evaluators in assessing dermatological parameters using digital photographs. This assessment was carried out by both an experienced Dermatologist and a group of trained Evaluators at NovoBliss Research. The purpose of involving the Dermatologist in the training was to establish a benchmark for evaluating inter-evaluator reliability between the Dermatologist and the trained evaluators. This process aimed to validate and qualify the evaluators as reliable assessors under the guidance of the Dermatologist.

METHODS

Study sites

This comprehensive non-interventional, training and validation study was conducted at NovoBliss Research Centre, Ahmedabad in the month of January 2023. The study encompassed evaluator training and photographic assessments aimed at establishing inter-evaluator reliability. The photographic evaluation utilized publicly accessible images from the internet.

Sample size

Given that this study did not entail any form of intervention, there was no enrolment of volunteers. Instead, the primary focus laid on training and validating the methodology employed for assessing various skin parameters. In this context, a total of 50 publicly accessible photographs from online sources were curated by the dermatologist for the purpose of conducting the evaluations.

Ethical conduct of the study

Some studies exempt from ethical review typically encompass research involving publicly available information, such as published biographies or newspaper reports.[4] Given the non-interventional nature of this training, which did not involve active human participation, and considering that the photographic evaluation utilized publicly available online images, formal ethical approval was not pursued for the training plan. It is noteworthy that the training procedures adhered to the principles outlined in ICH E6 (R2), which provides guidance on good clinical practice.

Study design

In this comprehensive non-interventional, training and validation study, the evaluators have assessed dermatological parameters of the photographs for the establishment of reliability for dermatologist-trained and validated evaluators. This involve evaluating different parameters that have been used during the dermatology study to ensure the consistency and accuracy of assessments made by these evaluators. The total duration of this comprehensive training including the training for evaluator and evaluation of photographs was of one day that included training on skin anatomy, explanation of dermatological parameters and photographic evaluations.

Procedure

The training plan was prepared and approved prior to the conduct of the training. The training encompassed various facets, beginning with a theoretical overview of skin anatomy presented via a comprehensive PowerPoint session. The Dermatologist selected 50 high-resolution photos per parameter for training. Both the Dermatologist and Trainee Evaluators used LED screens and computers to score these digital images. Stringent specifications and standardizations were followed, with separate scoring sheets used for each parameter per evaluator. Scoring took place at NovoBliss Research, and data were recorded and compiled into an Excel spreadsheet. Trained evaluators' scores were shared with a biostatistician for subsequent statistical analysis.

Training for evaluators

Ahmedabad based dermatologist with wide experience in the field of dermatology, conducted a training session at NovoBliss Research Centre, Ahmedabad in Jan 2023 on 5th Jan 23. The session covered skin anatomy through a PowerPoint presentation, followed by an explanation of dermatological parameters, including the Draize Scale for Irritation, Fitzpatrick Skin Type, Glogau Skin Age Classification, Physician Global Assessment using Griffith Scale, Investigator Global Assessment for Acne Severity, and Physician Global Assessment for brittle nail signs. A General Practicing Physician with 18 years of clinical and research experience, provided instruction on managing Adverse Events (AEs) and Serious Adverse Events (SAEs). The Director of Operations at NovoBliss Research with 18 years of clinical research experience, offered guidance on regulatory aspects and the need to establish inter-evaluator reliability between the Dermatologist and Dermatologist Trained Evaluators.

Photographic evaluation

The photographic evaluation method involved using various photographs assessing or evaluating the particular phenomenon or subject. This method is commonly used in the field such as anthropology, sociology, medicine, biology, horticulture, and dermatology where researcher use photographs to evaluate and analyse various aspect. For the training, the Dermatologist had selected the 50 high-resolution photographs for each parameter. Dermatologist and all Evaluators had used same single LED screen to view and score digital photographs using computer system. Only LED screen brightness for evaluation with screen resolution of 1920 x 1080, screen light level 90, and refresh rate 60p Hz was used for scoring of photographs along with maintained display scaling (Figure 1). Dermatological evaluations for irritation scoring using Draize Scale, Fitzpatrick Skin type, Glogau Skin age classification, IGA (Investigator Global Assessment) scoring for acne breakout and severity, PGA of facial skin appearance score using Griffith scale (0-8) for coarse and fine wrinkles, depigmentation (lentigines), elastosis, skin roughness, pore size, telangiectasia, sallowness, and skin laxity, and PGA for signs of brittle nails for Lamellar Onychoschizia, Ridging, Longitudinal splitting, Fragility/breakage, thickness, surface roughness, raggedness, and peeling was evaluated by dermatologist as well as the evaluator.

Draize scale was used for the evaluation of skin reactions namely erythema, edema, dryness and wrinkle on a severity scale of 0-4 where 0 score indicates no reaction and 4 score indicates severe erythema, edema, dryness, and wrinkles.[5] The Fitzpatrick skin type scale was used to classify skin types ranging from Type I to Type VI.[6] Glogau Skin Age was assessed to confirm the skin age type from 1 to 4.[7] To assess the acne severity and its improvement, one of the tools Investigator Global Assessment of acne (IGA) was used. The severity was assessed on the scale 0-4 where 0 score represents Clear skin with no inflammatory or noninflammatory lesions and 4 score represents severe, up to many noninflammatory lesions and may have some inflammatory lesions, but no more than a few nodular lesions.[8] Severity scoring of individual skin appearance parameters was scored on a severity scale of 0-9 where score 0 indicates No appearance, score 1-3 indicates mild, score 4-6 indicates moderate, and score 7-9 indicates severe. All participant's data was compiled into excel spreadsheet and scores of the trained evaluator were shared with biostatistician for statistical analysis. All the evaluators have passed the training satisfactorily, within the training plan acceptance criteria i.e. >80%, and are eligible for functioning as independent "Dermatologist Trained and Validated Evaluator" in any study conducted at NovoBliss Research.

Statistical analysis

The statistical analysis was executed using the R software (Version 4.3.1). To ascertain the correlation between the Evaluators and Dermatologist across all parameters, a Paired Correlation test was employed, with the exception of the Glogau Skin Age parameter. For the Glogau Skin Age, a Pearson Chi Square test was opted for, given the inherent characteristics of the data.

RESULTS

Total of 09 evaluators who were the paramedics staff were part of the training from NovoBliss Research, Gandhinagar, Gujarat, India.

Figure 1: Photographic evaluation.

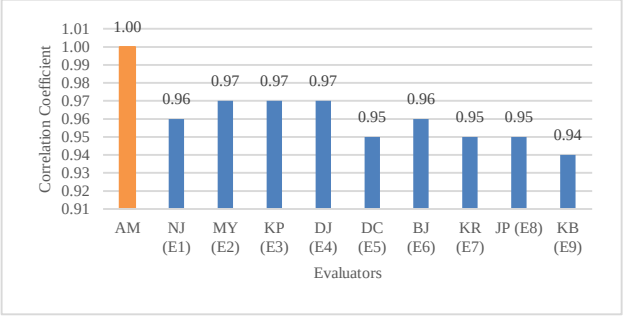

Figure 2: Correlation of evaluators with dermatologist for Fitzpatrick skin type.

Table 1: Photographic score's percentage agreement between evaluator and dermatologist.

Evaluator/Dermatologist	Correlation Value	% agreement	Acceptance criteria (%)
Evaluator 1	0.92	92	≥80
Evaluator 2	0.93	93	≥80
Evaluator 3	0.95	95	≥80
Evaluator 4	0.92	92	≥80
Evaluator 5	0.91	91	≥80
Evaluator 6	0.90	90	≥80
Evaluator 7	0.95	95	≥80
Evaluator 8	0.91	91	≥80
Evaluator 9	0.94	94	≥80

This training was completed in blinded fashion, and each evaluator passed the training satisfactorily, within the training plan acceptance criteria, and are eligible for functioning as independent "Dermatologist Trained and Validated Evaluator" in any study conducted at NovoBliss Research, based on the statistical analysis. Results of statistical analysis of photographic evaluation Inter-Evaluator reliability were released in January 2023.

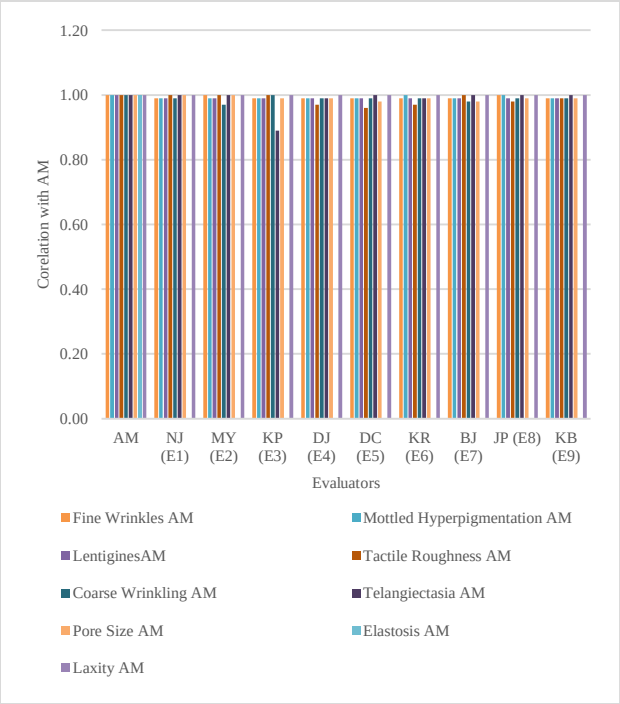

Figure 3: Correlation of evaluators with dermatologist for PGA using Griffith Scale for skin appearance.

On the basis of the data obtained it was found that the readings between the data set of Dermatologist and the data set of Trained Evaluator were highly corelated with each other. The evaluators who met the acceptance criteria of 80% as mentioned in the training plan among those evaluators and dermatologists were found suitable to perform dermatological assessments and photographic evaluation for studies that require such evaluations.

The theoretical training was done first followed by photographic evaluation training by the Dermatologist.

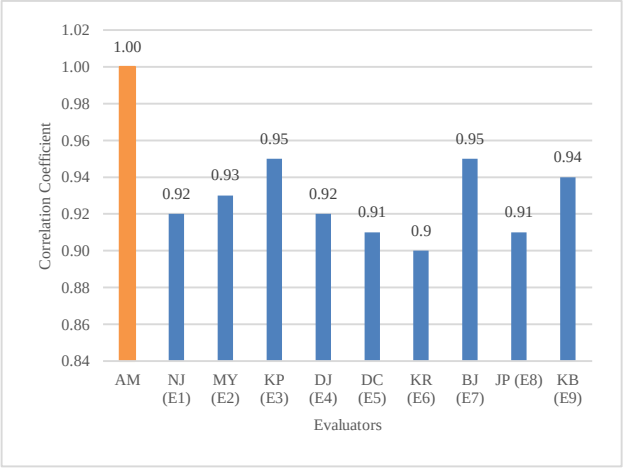

Figure 4: Correlation of evaluators with dermatologist for IGA scale for acne severity.

Table 2: Detailed Glogau skin age classification by dermatologist and evaluators.

Evaluators	X squared	Df	P value
Evaluator 1	87.00	9	<0.0001
Evaluator 2	114.83	9	<0.0001
Evaluator 3	89.05	9	<0.0001
Evaluator 4	103.50	9	<0.0001
Evaluator 5	104.99	9	<0.0001
Evaluator 6	109.79	9	<0.0001
Evaluator 7	132.20	9	<0.0001
Evaluator 8	101.31	9	<0.0001
Evaluator 9	119.87	9	<0.0001

Each scorer had done scoring using the same system under same lightening condition. Same lightening condition was checked using Luxmeter and recording done twice a day before start and end and observed same condition (Table 1). The Fitzpatrick skin type scale and Glogau Skin Age was assessed to confirm the skin age type. All the evaluators met the acceptance criteria for the parameter Fitzpatrick Skin Type (Figure 2). All the evaluators has ≥ 80% of score.

Table 3: Comparing dermatologist and evaluator scores on draize scale for skin irritation (erythema/dryness/ wrinkles and edema).

Evaluators	Correlation value with dermatologist (erythema/dryness/ wrinkles)	Correlation value with Dermatologist (Oedema)
Evaluator 1	0.96	0.98
Evaluator 2	0.96	0.96
Evaluator 3	0.97	0.95
Evaluator 4	0.97	0.94
Evaluator 5	0.97	0.91
Evaluator 6	0.98	0.96
Evaluator 7	0.98	0.93
Evaluator 8	0.97	0.94
Evaluator 9	0.97	0.94

Table 4: Correlation table for physical global assessment for brittle nails by dermatologist and evaluators.

Evaluators	Lamellar Onychoschizia	Ridging	Longitudinal Splitting	Fragility/ Breakage	Thickness	Surface Roughness	Raggedness	Peeling
Evaluator 1	0.97	0.99	0.95	0.98	1.00	0.95	0.96	1.00
Evaluator 2	0.97	0.98	1.00	0.97	1.00	0.97	0.95	0.97
Evaluator 3	0.96	0.97	0.82	0.97	1.00	0.94	0.96	0.97
Evaluator 4	0.96	0.98	0.95	0.98	1.00	0.95	0.93	0.95
Evaluator 5	0.98	0.96	0.71	0.97	1.00	0.95	0.92	0.95
Evaluator 6	0.97	0.97	0.87	0.98	1.00	0.94	0.96	0.97
Evaluator 7	0.99	0.97	0.91	0.99	1.00	0.93	0.92	0.98
Evaluator 8	0.97	0.98	0.78	0.98	1.00	0.97	0.94	0.91
Evaluator 9	0.97	0.98	0.83	0.97	1.00	0.98	0.95	1.00

Dermatologist/Dermatologist Trained Evaluator assessed the evaluation parameters of type I (No wrinkles) to Type IV (only wrinkles).

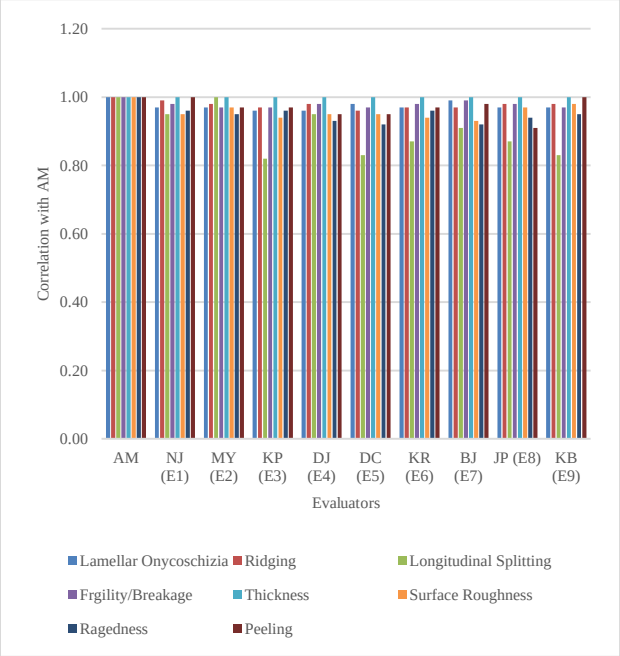

Figure 5: Correlation of evaluators with dermatologist for PGA for signs of brittleness of nails.

As per the statistical data obtained from pearson chi square test, the p value was found to be <0.05, which indicates the corelation between the data set of trained evaluators in comparison with data set of Dermatologist (Table 2). For PGA for sign of brittle nails, the parameters like lamellar onychoschizia, ridging, longitudinal splitting, fragility/breakage and thickness were scored on the scale of 0-5 where score 0 indicates none and score 5 indicates severe and parameters like roughness, raggedness, and peeling were scored on the scale of 0-3 where 0 score indicates none and 3 score indicates severe.[9] In the PGA Scale of Griffith's Scale (Figure 3), all the evaluator has passed and met the acceptance criteria for Fine Wrinkles, Mottled Hyperpigmentation, Lentigines, Tactile Roughness, Coarse Wrinkling, Telangiectasia, Pore Size, Elastosis, and Laxity with the severity score 0-9. In Irritation Scoring using Draize Scale, erythema, dryness, wrinkles on a 0point scale scored separately for each parameter and oedema on another 0-4 points scale as per the Draize Scale was assessed. All the evaluator met the acceptance criteria. All the evaluator has ≥80% of score (Table 3). Additionally, on the basis of the data obtained it was found that the readings between the data set of Dermatologist and the data set of Trained Evaluator were highly corelated with each other. Scoring sheets of evaluators were shared with the designated staff. Based on the statistical analysis for IGA for acne severity (Figure 4) which proved that there is highly correlation between readings of dermatologist and evaluators. Nail brittleness

is characterized by nails that split, flake, and crumble, become soft, and lose elasticity. The main clinical presentations of brittle nails (BN) that have been described are onychoschizia, onychorrhexis, superficial granulation of keratin, and worn-down nails. PGA for Nail Brittleness (Figure 5, Table 4) was scored by trained evaluators. In the PGA scale for nail brittleness, all the evaluator has passed and met the acceptance criteria for lamellar onychoschizia, ridging, longitudinal splitting, fragility/breakage. thickness, surface roughness, raggedness, and peeling. On the basis of the data obtained for above parameters was proven that the readings between the data set of dermatologist and the data set of trained evaluator were highly corelated with each other.

DISCUSSION

In clinical trials, the medication's effectiveness in treating various skin parameters like irritation, skin appearance, skin age, signs of brittle nail is determined using clinician-reported outcome measures, such as the Investigator Global Assessment (IGA) or Physician Global Assessment (PGA). Depending on the dermatologist's judgement, different outcomes are reported while evaluating these parameters. For clinical research in dermatology, there is a need to measure the relevant outcomes in the most accurate and consistent manner in frequent visits. To administer the assessment in the same way every single time and understanding the disease severity and scores in the same way by the evaluator becomes very crucial in such conditions.[10]

A clinical trial site must also have qualified Principal Investigators, study coordinators, and assessors, as well as suitable research protocols and recruitment tactics. Trained Evaluators aid and support Investigators with skin evaluation records, as well as monitor and report major adverse occurrences. Even when a dermatologist is unavailable, clinical readings are frequently missed, resulting in a loss of accurate data in a timely way with specified datasets. This study ensures the validity and reliability of all 9 evaluators.

The key benefit of this comprehensive evaluation training is that it can be utilised as a teaching tool in the future to collect accurate, timely readings from Dermatologists, Cosmetics, and Personal Healthcare Goods trial patients/subjects with the help of evaluators.[11] Dermatologist's Trained Evaluators plays a vital role in any clinical trial; however, this process does not come without challenges. Barriers that are associated with the application of the clinical trial results to one's own practice. An evaluator needs to carefully perform the assessment. For all patients, Evaluation should be a decision-making process for all patients, keeping the severity of their diseases in mind.[12] When using clinical trial study data, the one-size-fits-all approach is not always applicable.[13,15] The biological variability between patients is always there in spite of similar demographics and disease states needs to be taken into consideration.[14]

CONCLUSION

The dermatologist thoroughly explained each scoring scale and demonstrated how to assess conditions and assign scores using a diverse set of 50 images/parameters during the training. This approach aimed to enhance the reliability of dermatology clinical trial evaluations. The training was attended by nine evaluators, all of whom satisfactorily completed it. Statistical analysis revealed that each evaluator met the 80% acceptance criteria outlined in the training plan. These evaluators, alongside dermatologists, are deemed capable of performing dermatological assessments and photographic evaluations for relevant studies. They are eligible to function independently as "dermatologist trained and validated evaluators" in future research. Theoretical training equipped evaluators to understand various dermatological parameters, including draize scale, fitzpatrick skin type, glogau skin age classification, PGA with Griffith Scale, IGA (Acne), and PGA for brittle nail signs. Trained evaluators are now qualified to independently serve as "dermatologist validated and trained evaluators" in future studies. This approach is suitable for multi-centered, multi-visit dermatological clinical studies.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Apeksha Merja (MD, Dermatologist), Ms. Maheshvari Patel (PhD Scholar, Pharmacology), Dr. Nayan Patel (MBBS), Dr. Bhavna Nangha and Ms. Megha Yadav (Medical Writer) and Ms. Nistha Jani (MSc. Microbiology) for the kind collaboration and providing training. The authors are also thankful to the NovoBliss Research study team, evaluators for the training conduct, the statistical team for data analysis and the scientific writing team for support in the preparation of this manuscript.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: Not required

REFERENCES

1. Buranasirin P, Pongpirul K, Meephansan J. Development of a Global Subjective Skin Aging Assessment score from the perspective of dermatologists. BMC Res Notes. 2019;12(1):364.
2. Bhor U, Pande S. Scoring systems in dermatology. Indian J Dermatol Venereol Leprol. 2006;72:315-21.
3. Muhaba KA. Automatic skin disease diagnosis using deep learning from clinical image and patient information. Skin Health Dis. 2022;2:e81.
4. Does my study needs ethical approval, Cambridge Psychology Research Ethics Committee. Available at: https://www.bio.cam.ac.uk/psyres/approval. Accessed on 20 February 2023.
5. Draize J, Woodard G, Calevery H. Methods for the study of irritation and toxicity of substances applied

topically to the skin and mucous membranes. J Pharmacol Experiment Therap. 1944;2:377-90.
6. JI Y, Kimb BK, Suh DH. The validity and practicability of sun-reactive skin type I thorough VI. Arch Dermatol. 1988;5:869-71.
7. Glogau RG. Chemical peeling and aging skin. J Geriatr Dermatol. 1994;2:30-5.
8. Draft Guidance for Industry: Acne Vulgaris: Developing Drugs for Treatment. Clinical/Medical. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071292.pdf. Accessed on 20 February 2023.
9. Griffiths C, Wang TS, Voorhees TA. A photonumeric scale for the assessment of cuteneous photodamage. Arch Dermatol. 1998;12:347-51.
10. Sherber NS, Hoch AM, Coppola CA, Carter EL, Chang HL, Barsanti FR, Mackay-Wiggan JM. Efficacy and safety study of tazarotene cream 0.1% for the treatment of brittle nail syndrome. Cutis. 2011;87(2):96-103.
11. Anthony E, Kalpadakis-Smith A. Properly training physicians in dermatology trials. Applied Clin Trials. 2021;30:8.
12. Green L. Explaining the role of the nurse in clinical trials. Nurs Stand. 2011;25:35-9.
13. Fahey T. Applying the results of clinical trials to patients to general practice: perceived problems, strengths, assumptions, and challenges for the future. Br J Gen Pract. 1998;48:1173-8.
14. Iwashyna TJ, Deane AM. Individualizing endpoints in randomized clinical trials to better inform individual patient care: the TARGET proposal. Crit Care. 2016; 20:218-25.
15. Zarbin MA. Challenges in applying the results of clinical trials to clinical practice. JAMA Ophthalmol. 2016;134:928-33.

Cite this article as: Merja AM, Patel NK, Patel MN, Yadav MK, Jani NA. Establishing reliability of dermatologist trained and validated evaluators for assessments of dermatological parameters: a comprehensive overview of concepts and techniques. Int J Res Dermatol 2023;9:262-8.