Open access • Journal Article • DOI:10.1111/ICS.12563

An automatic procedure that grades some facial skin structural signs: agreements and validation with clinical assessments made by dermatologists. — Source link

Frederic Flament, Maja A. Hofmann, Elia Roo, Catherine Raimbault-Gerard ...+9 more authors

Institutions: L'Oréal, Charité, Rafael Advanced Defense Systems, Metz

Published on: 01 Oct 2019 - International Journal of Cosmetic Science (Int J Cosmet Sci)

An automatic procedure that grades some facial skin structural signs

agreements and validation with clinical assessments made by dermatologists

Flament, F.; Hofmann, M.; Roo, E.; Raimbault-Gerard, C.; Chuberre, B.; Zhang, Y.; Jiang, R.; Kezele, I.; Zhang, J.; Elmozino, E.; Coquide, J.; Arcin, V.; Aarabi, P.

Published in:
International Journal of Cosmetic Science

DOI:
10.1111/ics.12563

Publication date:
2019

Document version:
Accepted manuscript

Citation for pulished version (APA):
Flament, F., Hofmann, M., Roo, E., Raimbault-Gerard, C., Chuberre, B., Zhang, Y., Jiang, R., Kezele, I., Zhang, J., Elmozino, E., Coquide, J., Arcin, V., & Aarabi, P. (2019). An automatic procedure that grades some facial skin structural signs: agreements and validation with clinical assessments made by dermatologists. *International Journal of Cosmetic Science*, *41*(5), 472-478. https://doi.org/10.1111/ics.12563

MR. FREDERIC FLAMENT (Orcid ID : 0000-0003-1128-9535)

Article type : Original Article

Corresponding author mail id: fflament@rd.loreal.com

An automatic procedure that grades some facial skin structural signs: agreements and validation with clinical assessments made by dermatologists.

FLAMENT Frederic[1], HOFMANN Maja[2,3], ROO Elia[4], RAIMBAULT-GERARD Catherine[5], CHUBERRE Bertrand[6], ZHANG Yuze[7], JIANG Ruowei[7], KEZELE Irina[7], ZHANG Jingyi[7], ELMOZINO Eric[7], COQUIDE Jerome[8], ARCIN Vincent[8], AARABI Parham[7]

[1] L'Oréal Research and Innovation, Clichy, France.

[2] Dermatology, venereology and allergy department, Charité Medicine University, Berlin, Germany.

[3] Department of dermatology, University Odense, Denmark.

[4] Private Dermatology Consultancy Practice, Madrid, Spain.

[5] Private Dermatology Consultancy Practice, Metz, France.

[6] Active Cosmetics International, Global Medical Relations and Communications, Levallois, France.

[7] ModiFace – A L'Oréal Group Company, Toronto, Canada

[8] L'Oréal CDO – Digital Service Factory, Clichy, France.

Abstract

Objective: To confirm the robustness and validity of an automatic scoring system, algorithm-based, that grades the severity of nine facial signs through "selfies" smartphones pictures taken by European Caucasian women through dermatological assessments.

Methods: 157 Caucasian women from three countries (France, Germany, Spain), of different ages (20–75 years), took one "selfie" image by the frontal camera of their smartphones whereas local dermatologists photographed them with the back camera of the same smartphone. The same nine facial signs of these subjects were initially graded by these local dermatologists, using referential Skin Aging Atlases. All 314 "selfies" images were then further automatically analyzed by the algorithm. The severity of facial signs (wrinkles, pigmentation, ptosis, skin folds etc.) were statistically compared to the assessments made by the three dermatologists, taken as ground truth.

Results: Highly significant coefficients of correlation ($p < 0.001$) were found in the three cohorts between the grades provided by the system and those from dermatologists in live. The back camera – of a better resolution than the frontal one – seems affording slightly more significant correlations. However, although significantly correlated, the signs of vascular disorders and cheek skin pores present some disparities that are likely linked to the technical diversity of smartphones or self-shootings, leading to lower coefficients of correlations.

Conclusion: This automatic scoring system offers a promising approach in the harmonization of Dermatological assessments of skin facial signs and their changes with age or the follow up of anti-aging treatments.

Resumé

Objectif : De confirmer la validité et la solidité d'un système de scorage automatique qui quantifie la sévérité de neuf signes du visage à partir de photographies de type "selfies" prises par des femmes Caucasiennes Européennes d'âge différents.

Méthodes : 157 femmes Caucasiennes de trois pays différents (France, Allemagne, Espagne), d'âges différents (20–75 ans) ont pris un « selfie » avec la caméra frontale de leur téléphone tandis que le dermatologue local les a photographiées à l'aide de la caméra dorsale du même appareil. Les neuf signes faciaux ciblés par le système de scorage automatique ont été préalablement évalués par trois dermatologues locaux, utilisant des Atlas référentiels du vieillissement cutané. Les 314 images obtenues furent ensuite analysées automatiquement par l'algorithme. Les sévérités des neuf signes (rides, ptose, plis, pigmentation...) ont été ensuite comparées à celles établies par les dermatologues, considérées comme références absolues.

Résultats : De très significatifs coefficients de corrélation ($p < 0,001$) ont été trouvés dans les trois cohortes entre les scores fournis par le système et ceux issus des évaluations des dermatologues des visages durant la visite des volontaires. La caméra du dos des smartphones – de meilleure résolution que la frontale – semble fournir de légèrement meilleures significativités. Cependant, bien que significativement corrélés, les signes des désordres vasculaires et des pores cutanés des joues

montrent quelques disparités, dues possiblement à la diversité technique des smartphones ou celle des prises de vue, conduisant à de plus faibles coefficients de corrélation.

Conclusion : Ce système de quantification automatique semble offrir une approche prometteuse dans l'harmonisation des évaluations dermatologiques des signes faciaux et leurs modifications liées à l'âge et/ou le suivi de traitements à vocation antivieillissement cutané.

Introduction

The human face affronts the combined assaults of a dictatorial aging process and those from external foes such as exposures to sun [1] and/or aerial pollution [2–4]. These lead to a progressive alteration of some facial traits, being structural (wrinkles, ptosis etc.) or pigmentary, the changes of which are now well documented and clinically evaluated [5–13].

In a previous paper [14], an algorithm was presented, specifically designed to automatically detecting facial signs (wrinkles, pigmentary disorders, ptosis, vascular disorders etc.) and grading their clinical severities, from selfies images/smartphones. The latter were taken by 1140 women of three different ethnics and ages, under variable conditions (lighting, distance, facial expressions etc.). This first study [14] showed that smartphones of an optical memory below 3 Megapixels led to less reproducible results. Comparisons with expert's assessments, who graded the various severities using referential skin aging atlases [15–19], were shown in significant agreement in most cases. The system was found weakly influenced by lighting conditions or facial expressions and leads to global and significant agreements with experts' assessments – each subject or photograph was evaluated by a quorum of 12 experts.

These promising results nevertheless needed another challenging issue. To which extent this automatic scoring system would conform with assessments from dermatologists became a mandatory proof of truth. Accordingly, a study involving 157 Caucasian women, differently aged (20–75 years) from three European countries (France, Germany, Spain) was organized under comparable conditions to the previous investigation [14], supervised by three local dermatologists – each subject being observed by only one dermatologist. The results of this study are the foci of the present paper.

Material and Methods

Subjects

157 Caucasian European born women (French, German, Spanish) differently aged (20–75 years) and residing in Metz, Berlin and Madrid for more than 15 years, participated to the study. Table I

illustrates the respective distributions of subjects by countries and ages, after being recruited by local agencies dedicated to part time occupations. As prerequisite, this recruitment phase made sure that each subject was: i) used to take “selfies” and ii) that her smartphone presented two cameras (front and back) of at least 3 Megapixels capacity. They were enrolled by the local dermatologist who clinically checked the absence of a particular skin disease or disorder (vitiligo, rosacea, acne, melasma etc.) or permanent facial make-up. All women were asked to attend with a bare facial skin, i.e. no application of make up or skin care products a day before. All were fully informed, in their respective languages, on the objective of the study and signed an informed consent that mentioned that all individual images, once analyzed, will be deleted.

Protocol (illustrated in Figure 1)

Step1: In each country, the local dermatologist (CR-G, MH, ER) graded the same 9 facial signs, using three referential Skin Aging Atlases [15,16,19]. These signs, and their respective grading scales, are illustrated in table II.

Step 2: All subjects were asked to take one “selfie” (full face, frontal camera) in presence of the dermatologist (same lighting conditions: indoor artificial diffuse light). All subjects were asked to adopt the most neutral expression when taking this “selfie”.

Step 3: Dermatologist used the same smartphone to take a full-face “standardized” picture of the subject using the back camera (of a higher resolution than the frontal camera) in the exact same lighting conditions as the “selfie” adopting the same neutral expression.

Step 4: Both images were then automatically analyzed by the developed algorithm used in the previous study [14] and the resulting automatic gradings of the 9 facial signs were sent to our secured website under blind codes such as FR8, GE11 or SP17.

Step 5: Statistical analysis of the data afforded by the 2 automatic gradings (front and back cameras), taking those provided by the dermatologists under live conditions as ground truth/control.

Statistics

The on-live dermatological grading of each subject was used as the ground truth reference to assess the accuracy of our automatic scoring system, in addition to the previous experiments realized with experts [14]. Predicted scores for both frontal "selfies" and back camera images were compared to this ground truth score, per subject. The following calculations were used to evaluate the automatic scoring system:

i) Median absolute error (MAE), as the median of the absolute differences between the predicted and the ground truth scores across all samples, applied separately to frontal “selfies” and

"standardized" back camera images. On each subject, the calculation of the MAE applied to all the 9 clinical facial signs.

The median metric for robustness was chosen against any potential bias in ground truth scores due to "human error", as the ground truth scores are defined by one dermatologist per subject.

ii) The fraction of samples for which the absolute error is below a given threshold T of grading score was calculated. Three different thresholds were used (T: 0.5, 1 and 1.5).

Further, the correlation of score trends between the ground truth expert scores and the scores predicted by our automatic scoring system was analyzed through Spearman Correlation Coefficients Tests.

All inter-correlations were carried out through Student t test, taking a $p < 0.05$ as threshold of significance, using the SPSS software package (SPSS/IBM, Chicago/ILL, USA).

For all the 157 subjects, 9 clinical facial signs were assessed using 3 different methods.

For each clinical sign, the correlations between each method (2 pairs of methods to be compared) and their significance were calculated using Spearman Coefficient Correlation Tests i) for each country (around 50 models / country) and ii) overall, taking all models into account (N = 157).

Results

Table III gives the consolidated paired comparisons obtained between automatic grading's from frontal and back cameras and those obtained from the dermatologists, for the 157 studied women. The correlation coefficient (r values) were found highly significant and systematically (albeit slightly) higher with a back camera than the frontal one, of a better resolution. However, although significant, the coefficients of correlations for Vascular disorders and Cheek skin pores appear of much lower values than those of the 7 other facial signs.

Tables IV, V and VI give the respective values, country by country.

i) French cohort

ii) German cohort

iii) Spanish cohort

The MAEs and percentages of population samples for which the error is below than a given threshold per sign are given in tables VII and VIII. The results are shown separately for frontal pictures (Table VII) and "standardized" back photographs (Table VIII). In these two tables, results are both presented globally (the three cohorts) and as separate (i.e. per country of origin). Overall, the majority of signs

displays a MAE well beyond 1, and the percentage of the population with absolute error less than 1.5 is around 80 %. Median absolute errors average about 0.8 grade unit and, for some local panels, fall below 0.55 grade unit.

Discussion

The results of this multi centric study organized in three European countries, under the supervisions of three local dermatologists, globally confirm the robustness of the dedicated algorithm versus an on-live ground truth. These results globally confirmed those previously found by expert's assessments on pictures and "selfies" [14]. A few points have however to be commented and discussed. At first, all data resulted from different models of smartphones and three different local lighting conditions. Second, when analyzing the global results (N = 157), the optical superiority of the back camera over the frontal one is unsurprising, although their differences (in terms of r values) are indeed close. When analyzing results (the correlation coefficients r) country by country, such optical superiority appears indeed less obvious as, in many cases, both cameras led to comparable coefficients of correlations, making the use of a frontal camera of a $\geq$ 3 Megapixels resolution an already highly satisfying criterion. At this stage of comments, it is likely that some variable (and unavoidable) factors inherent to smartphones (size of the screens, more or less efficient optical components etc.) or to the "selfie" shootings (un-precise framing, different positioning etc.) play a major role. Third, highly significant coefficients of correlation obtained on 7 signs out of 9 indicate that the automatic scoring system applies well in different conditions and subjects, in close agreements with the dermatologist's assessments. The cases of vascular disorders or cheek skin pores appear less clear. Although significant, their respective r values are often much lower than the 7 other signs in the three cohorts, mostly among Spanish women. Of note, vascular disorders were found well correlated in the French and German cohorts. It cannot be excluded that the different innate facial skin complexions (fair, dark) modulate the contrast with these reddish elements, i.e. less contrasted in darker skins (e.g. higher photo-types) than in fairer complexions. On such aspect, it can be reasonably assumed that Spanish women may present a slightly darker tone, more susceptible to be frequently sun-exposed than French or German women of northerner latitudes. In addition, vascular disorders may slightly vary with heat, emotions (flushing, blushing). Combined, these two major aspects make vascular disorders more difficult to be assessed with accuracy.

The detection and analysis of cheek skin pores in all three cohorts seem, as for now, weak points of the automatic scoring system, calling for a deeper analysis and possible improvements of the algorithm especially regarding the limits of resolution of "selfies", towards an accurate grading. Of

note, skin pores are particularly of a difficult assessment by naked eyes, with regard to their variable dimensions, depths and densities. A previous work, using a magnifying electronic instrument showed how these skin features are age-dependent within a same ethnic [20].

Whatsoever, the present study confirms the interest of an automatic scoring system on "selfies" smartphones' images that collects and quantifies the clinical severities of some facial signs in close agreement with dermatologists or experts. The fact that each subject was observed by only one dermatologist whereas in the creation and validation stage [14], each photograph was evaluated by a quorum of 12 experts explains the slight differences in accuracy. The agreements between the three dermatologists could not obviously be strictly compared as each of them examined different subjects. However, analyzing the respective correlations (in the three countries, Tables IV, V and VI) between their own grading's and those provided by the automatic system, led to non-significant slightly lower values (r decreased by about 0.06 in average) to those obtained by an automatic grading. A comparable result was obtained with regard slightly diverging assessments between experts [14]. In brief, the automatic system seems offering a slightly better accuracy than the one given by an unavoidable individual subjectivity of the human vision.

As for the dermatological domain, this automatic scoring system first represents a promising tool to collect harmonized data, storing records of the facial status of patients when subjected to different treatments, in total confidentiality. Hence, this automatic scoring system appears a valuable approach to implement the knowledge and expertise of the dermatological community, through very easy and common procedures, now daily used by millions of people, worldwide. Last but not least, this system should greatly facilitate the communication between MD's and their patients as *"An image is worth a thousand words"*.

Acknowledgments

Authors want to thank Mr Didier Saint-Leger, Mrs Laurie Jacquet, Mrs Hajar Ksikes, Mr Christian Bouchard, Mrs Myriam Bekkar-Schneider, Mrs Brigitte Liberman, Mrs Pascale Mora, Mrs Isabelle Walter, Mrs Cecile Jolly, Mr Stephan Garandet, Mrs Esohe Omoyuri, Mrs Camille Kroely, Mrs Sarah Koehler, Mrs Brittany Wanka, Mrs Jennifer Tidy, Mrs Lauren Sarda Dutilh, Mrs Seema Dwivedi, Mrs Diana Marcela Delgado Cortes, Mrs Chloe Fatou, Mrs Aude Charbonneau and Mr Thierry Lageat not only for their great help in completion of this paper and global study but also for their strong support and enthusiasm.

Conflict of interest

All authors are employees of L'Oréal Research & Innovation Dpt and its subsidiary Modiface company.

Funding

The costings of the present work were met by Modiface (a L'Oréal affiliate).

References

1. Gilchrest BA. Skin aging and photo aging: an overview. *J Am Acad Dermatol.* **21**, 610–613 (1989).
2. Krutmann J, Bouloc A, Sore G, Bernard BA, Passeron T. The skin aging exposome. *J Derm Sci.* **85 (3)**: 152–161 (2017).
3. Krutmann J, Liu W, Li L, Pan X, Crawford M, Sore G, Seite S. Pollution and skin: from epidemiological and mechanistic studies to clinical implications. *J Derm Sci.* **76 (3)**: 163–168 (2014).
4. Flament F, Bourokba N, Nouveau S, Li J, Charbonneau A. A severe chronic outdoor pollution alters some facial signs in Chinese women. *Int J Cosmet Sci.* **40(5)**: 467–481 (2018).
5. Levakov A, Vuckovic N, Doali M et al. Age-related skin changes. *Med Pregl.* **65(5-6)**: 191–195 (2012).
6. Mayes, A.E., Murray, P.G., Gunn, D.A. et al. Environmental and lifestyle factors associated with perceived facial age in Chinese women. PLoS One 5, e15270 (2010).
7. Flament F, Bazin R, Laquieze S, Rubert V, Simonpietri E, Piot B. Effect of the sun on visible clinical signs of aging in Caucasian skin. *Clin Cosmet Investig Dermatol.* **6**: 221–232 (2013).
8. Flament F, Bazin R, Qiu H, Ye C, Laquieze S, Rubert V, Decroux A, Simonpietri, Piot B. Solar exposure(s) and facial clinical signs of aging in Chinese women: impacts upon age perception. *Clin Cosmet Investig Dermatol.* **10**: 75–84 (2015).
9. Flament F, Amar D, Feltin C, Bazin R. Evaluating age-related changes of some facial signs among men of four different ethnic groups. *Int J Cosmet Sci.* **40**: 502–515 (2018).
10. Akiba S, Shinkura R, Miyamoto K, Hillebrand G, Yamaguchi N, Ichihashi M. Influence of chronic UV-exposure and lifestyle on facial skin photoaging – results from a pilot study. *J Epidemiol.* **9(6 Suppl)**: S136–142 (1999).
11. Nkengne A, Bertin C. Aging and facial changes-documenting clinical signs, part 1: clinical changes of the aging face. *Skinmed.* **11(5)**: 281–286 (2013).

12. Flament F, Qiu H. Effects of seasonality and a daily photo-protection upon some facial signs of Chinese women. *Int J Cosmet Sci.* **39**, 256–268 (2017).
13. Flament F, Gautier B, Benize AM, Charbonneau A, Cassier M. Seasonally-induced alterations of some facial signs in Caucasian women and their changes induced by a daily application of a photoprotective product. *Int J Cosmet Sci.* **39**, 664–675 (2017).
14. Jiang R, Kezele I, Flament F et al. A new procedure, free from human assessment, that automatically grades some facial skin structural signs. Comparison with assessments by experts, using referential atlases of skin ageing. *Int J Cosmet Sci.* **41**: 67–78 (2019).
15. Bazin R, Doublet E. Skin Aging Atlas. Volume 1, Caucasian Type. Paris: Editions Med'Com, (2007).
16. Bazin R, Flament F. Skin Aging Atlas. Volume 2, Asian Type. Paris: Editions Med'Com, (2010).
17. Bazin R, Flament F, Giron F. Skin Aging Atlas. Volume 3, Afro-American Type. Paris: Editions Med'Com, (2012).
18. Bazin R, Flament F, Rubert V. Skin Aging Atlas. Volume 4, Indian Type. Paris: Editions Med'Com, (2015).
19. Flament F, Bazin R, Qiu H. Skin Aging Atlas. Volume 5, Photo-aging Face & Body. Paris: Editions Med'Com, (2017).
20. Shaiek A, Flament F, François G, Lefebvre-Descamps V, Barla C, Vicic M, Giron F, Bazin R. A new tool to quantify the geometrical characteristics of facial skin pores. Changes with age and a make-up procedure in Caucasian women. *Skin Res Technol.* **23**(2), 249–257 (2017).

Panels, Location / Age clusters	Panel 1 (Spain)	Panel 2 (Germany)	Panel 3 (France)	Total
20–30y	10	8	4	22
31–40y	9	10	8	27
41–50y	10	10	13	33
51–60y	14	12	17	43
61–75y	7	11	14	32
Total	50	51	56	157

Table I: Distribution of subjects photographed and scored by dermatologists according to locations and age-groups.

Facial Signs	Definition of scored observation	Scale	Visual
Forehead wrinkles	Depth of the transverse wrinkles on the forehead.	0-5	
Glabellar wrinkles	Depth of vertical wrinkles between eyebrows.	0-5	
Periorbital wrinkles (upper cheek area)	Depth of folds at malar area below Crow's feet, eye orbit excepted.	0-5	
Nasolabial fold	Depth of the fold present on face between the base of the nose and lips.	0-5	
Marionette lines	Depth of folds at the corner of lips.	0-6	

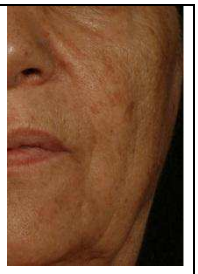
Density of pigmentary spots	Number of spots per area unit on the cheek area according to grades of Atlas.	0-5	
Ptosis of lower part of the face	Sagging severity of the lower parts each side of the chin.	0-5	
Vascular disorders	Diffused redness and microvessels visible on the face and especially cheeks.	0-7	
Cheek skin pores	Size of visible pores on the cheek irrespective of their densities.	0–5	

Table II: The nine Clinical facial signs of Caucasian women, related to Wrinkles & skin texture, Sagging and lack of Firmness, Pigmentation disorders, Cheek visible pores or Vascular disorders.

Facial signs / Correlations	Live dermatologist score Vs. Frontal "Selfies" analyzed by automatic clinical scoring system	Live dermatologist score Vs. Back Camera analyzed by automatic clinical scoring system
Forehead wrinkles	0.85 $p < 0.001$	0.89 $p < 0.001$
Glabellar wrinkles	0.86 $p < 0.001$	0.90 $p < 0.001$
Periorbital wrinkles (upper cheek area)	0.75 $p < 0.001$	0.78 $p < 0.001$
Nasolabial fold	0.82 $p < 0.001$	0.88 $p < 0.001$
Marionette lines	0.85 $p < 0.001$	0.90 $p < 0.001$
Density of pigmentary spots	0.67 $p < 0.001$	0.71 $p < 0.001$
Ptosis of lower part of the face	0.84 $p < 0.001$	0.88 $p < 0.001$

Vascular disorders	0.32 p < 0.001	0.40 p < 0.001
Cheek skin pores	0.28 p < 0.001	0.32 p < 0.001

Table III: Correlation coefficients values (r) of the total cohort of European women (N=157), and their respective significances, between live gradings by dermatologists and those obtained by the automatic clinical scoring system obtained either by smartphones' frontal or back cameras of their smartphones.

Facial signs / Correlations	**Live dermatologist score Vs. Frontal "Selfies" analyzed by automatic clinical scoring system**	**Live dermatologist score Vs. Back Camera analyzed by automatic clinical scoring system**
Forehead wrinkles	0.86 p < 0.001	0.92 p < 0.001
Glabellar wrinkles	0.84 p < 0.001	0.92 p < 0.001
Periorbital wrinkles (upper cheek area)	0.78 p < 0.001	0.88 p < 0.001
Nasolabial fold	0.82 p < 0.001	0.94 p < 0.001
Marionette lines	0.80 p < 0.001	0.91 p < 0.001
Density of pigmentary spots	0.55 p < 0.001	0.59 p < 0.001
Ptosis of lower part of the face	0.77 p < 0.001	0.86 p < 0.001
Vascular disorders	0.62 p < 0.001	0.72 p < 0.001
Cheek skin pores	0.39 p = 0.003	0.55 p < 0.001

Table IV: French cohort: correlations coefficients and their respective significance, between live scoring by the French dermatologist and results obtained by the automatic clinical scoring system provided by smartphones frontal or back cameras.

Facial signs / Correlations	Live dermatologist score Vs. Frontal "Selfies" analyzed by automatic clinical scoring system	Live dermatologist score Vs. Back Camera analyzed by automatic clinical scoring system
Forehead wrinkles	0.84 p < 0.001	0.83 p < 0.001
Glabellar wrinkles	0.88 p < 0.001	0.90 p < 0.001
Periorbital wrinkles (upper cheek area)	0.73 p < 0.001	0.76 p < 0.001
Nasolabial fold	0.87 p < 0.001	0.88 p < 0.001
Marionette lines	0.90 p < 0.001	0.90 p < 0.001
Density of pigmentary spots	0.72 p < 0.001	0.79 p < 0.001
Ptosis of lower part of the face	0.91 p < 0.001	0.89 p < 0.001
Vascular disorders	0.79 p < 0.001	0,77 p < 0.001
Cheek skin pores	0.18 p = 0.218	0.25 p = 0.078

Table V: German cohort: correlations coefficients and their respective significance, between live scoring by the German dermatologist and results obtained by automatic clinical scoring system provided by smartphones frontal or back cameras.

Facial signs / Correlations	Live dermatologist score Vs. "Selfies" analyzed by automatic clinical scoring system	Live dermatologist score Vs. Back Camera analyzed by automatic clinical scoring system
Forehead wrinkles	0.90 p < 0.001	0.90 p < 0.001
Glabellar wrinkles	0.91 p < 0.001	0.92 p < 0.001
Periorbital wrinkles (upper cheek area)	0.80 p < 0.001	0.77 p < 0.001
Nasolabial fold	0.89 p < 0,001	0.89 p < 0.001
Marionette lines	0,85	0.89

	p < 0.001	p < 0.001
Density of pigmentary spots	0.80 p < 0.001	0.80 p < 0.001
Ptosis of lower part of the face	0.91 p < 0.001	0.88 p < 0.001
Vascular disorders	0.34 p = 0.014	0.39 p = 0.005
Cheek skin pores	0.44 p = 0.001	0.40 p = 0.004

Table VI: Spanish cohort: coefficients of correlation and their respective significance, between live scoring by the Spanish dermatologist and results obtained by the automatic clinical scoring system provided by smartphones (frontal or back cameras).

	Forehead wrinkles	**Glabellar wrinkles**	**Periorbital wrinkles**	**Nasolabial fold**	**Marionette lines**	**Pigmentary disorders**	**Ptosis of lower face**	**Vascular disorders**	**Cheek skin pores**	**Average**
Global population										
MAE	0.56	0.63	0.89	0.68	0.63	1.21	0.48	0.81	1.45	0.82
%MAE<0.5	46	37	31	39	39	23	44	33	19	35
%MAE<1.0	76	77	56	65	65	39	63	50	30	58
%MAE<1.5	91	93	71	88	84	54	82	66	54	76
France										
MAE	0.85	0.64	1.27	1.28	0.79	1.46	0.89	1.94	2.18	1.26
%MAE<0.5	30	34	20	21	38	18	38	4	4	23
%MAE<1.0	70	77	46	45	59	29	51	9	5	43
%MAE<1.5	88	89	61	68	82	52	65	27	21	61
Germany										
MAE	0.36	0.53	0.98	0.57	0.46	0.72	0.31	0.21	0.94	0.56
%MAE<0.5	57	40	28	46	46	36	53	45	36	43
%MAE<1.0	80	81	51	76	71	54	78	79	57	70
%MAE<1.5	91	95	66	100	85	59	97	97	87	86
Spain										

MAE	0.37	0.63	0.53	0.45	0.75	1.37	0.47	0.33	1.43	0.70
%MAE<0.5	54	38	48	54	36	20	44	58	20	41
%MAE<1.0	80	74	72	78	68	38	65	78	32	65
%MAE<1.5	94	94	86	98	84	52	91	92	58	83

Table VII: Values of the Median Absolute Errors (MAE) in the scores of 9 facial signs obtained by the automatic scoring system on "selfies" pictures (frontal camera) and their changes with different selected thresholds.

	Forehead wrinkles	**Glabellar wrinkles**	**Periorbital wrinkles**	**Nasolabial fold**	**Marionette lines**	**Pigmentary disorders**	**Ptosis of lower face**	**Vascular disorders**	**Cheek skin pores**	**Average**
Global population										
MAE	0.47	0.63	0.91	0.65	0.65	1.27	0.74	0.77	1.44	0.84
%MAE<0.5	50	39	32	38	37	27	33	34	18	34
%MAE<1.0	75	73	54	71	62	40	56	50	29	57
%MAE<1.5	93	89	71	88	79	56	73	65	53	74
France										
MAE	0.79	0.76	1.11	0.99	0.69	1.38	1.12	2.04	2.19	1.23
%MAE<0.5	32	34	20	27	38	21	29	4	2	23
%MAE<1.0	66	71	43	52	64	34	45	7	7	43
%MAE<1.5	93	84	63	73	75	55	71	27	21	62
Germany										
MAE	0.29	0.46	1.02	0.56	0.44	0.60	0.35	0.20	1.00	0.55
%MAE<0.5	67	50	25	43	45	35	43	40	38	43
%MAE<1.0	82	77	48	81	62	50	62	83	50	66
%MAE<1.5	91	93	67	98	79	63	70	93	85	82
Spain										
MAE	0.45	0.70	0.47	0.53	0.82	1.38	0.66	0.35	1.36	0.75
%MAE<0.5	56	34	54	46	30	26	30	64	18	40
%MAE<1.0	78	70	72	82	60	38	63	78	34	64
%MAE<1.5	96	92	86	94	84	52	79	90	58	81

Table VIII: Values of the Median Absolute Errors (MAE) in the scores of 9 facial signs obtained by the automatic scoring system on “standardized” back camera pictures and their changes with different selected thresholds.

Frontal "Selfies"	Back Camera Picture	Facial signs	Live derm score	Frontal "selfie" scored by AI	Back camera scored by AI
		Forehead wrinkles	1.20	1.13	0.94
		Glabellar wrinkles	2.40	1.87	1.81
		Periorbital wrinkles	1.80	1.42	1.37
		Nasolabial fold	2.20	2.33	2.39
		Marionette lines	1.80	2.51	1.93
		Density of spots	3.80	2.96	2.54
		Ptosis of lower face	1.20	1.19	1.09
		Vascular disorders	2.20	2.05	2.11
		Cheek skin pores	2.60	1.09	1.08

Figure 1: Description of all data and images collection and their subsequent analysis on a 42 years old Spanish woman.