ORIGINAL ARTICLE

Validation of an Optical Pencil Method to Estimate the Affected Body Surface Area in Psoriasis ☆

A. Reolid,[a,*] O. Servitje,[b] M. Ginarte,[c] A. Pérez-Ferriols,[d] A. Vélez,[e] M. Aragües,[a] M. Pereiro,[c] J.L. Sánchez-Carazo,[d] J. Garrido,[f] E. Daudén[a], Grupo de Trabajo de evaluación del Body Surface Area (BSA)

[a] Servicio de Dermatología, Hospital Universitario de la Princesa, Instituto de Investigación Sanitaria la Princesa (IIS-IP), Madrid, Spain
[b] Servicio de Dermatología, Hospital de Bellvitge, Barcelona, Spain
[c] Servicio de Dermatología, Complexo Hospitalario Universitario, Santiago de Compostela, Spain
[d] Servicio de Dermatología, Hospital General de Valencia, Valencia, Spain
[e] Servicio de Dermatología, Hospital Reina Sofía, Córdoba, Spain
[f] Departamento de Psicología Social y Metodología, Facultad de Psicología, Universidad Autónoma de Madrid, Madrid, Spain

Abstract
Background: Body surface area (BSA) affected by psoriasis is one of the most often used measures for assessing severity, but this method has shortcomings.
Objective: To validate a new way to estimate BSA.
Material and method: Prospective, multicenter study in 56 patients with psoriasis. Each patient was evaluated by 2 dermatologists in 2 visits to the same hospital. Each dermatologist used 2 methods for estimating BSA: the traditional visual estimation in which the area of the palm equals 1% of the total body surface and an optical pencil (OP) method in which the affected area is drawn on a touch screen. Software in the application then calculates the BSA.
Results: Overall concordance between the 2 methods was acceptable according to an intraclass correlation coefficient (ICC) of 0.87. However, the limits of agreement were unacceptably large and there was systematic bias: traditional estimates were consistently greater than OP calculations. Concordance between the methods was better (ICC > 0.8) on the trunk and lower extremities. Intraobserver reliability was excellent with both methods (ICCs, 0.97 and 0.98 for the traditional and OP estimates, respectively). Interobserver reliability was also high (ICCs, 0.91 and 0.94 for the traditional and OP methods), although the mean BSA differed significantly between observers. The ICCs were much lower for BSA estimates on the head.

Conclusions: This study to validate the OP method for estimating the affected BSA in patients with psoriasis shows good agreement between the OP and traditional approaches. The OP calculations also showed less variance and better interobserver reliability.

Validación del método de cuantificación del área corporal afectada por la psoriasis mediante lápiz óptico

Resumen
Introducción: La determinación de la superficie corporal afectada, ''*Body Surface Area*'' (BSA), es una de las escalas de medida más empleadas en la evaluación de la gravedad de la psoriasis, pero no está exenta de inconvenientes.
Objetivo: Validación de un nuevo sistema de medida del BSA.
Material y método: Estudio multicéntrico, prospectivo, que incluyó 56 pacientes con psoriasis. Cada paciente fue evaluado en dos visitas por dos dermatólogos del mismo Centro que valoraron BSA mediante dos procedimientos: método visual ''tradicional'' (MT), palma mano = 1%, y; el método ''lápiz óptico'' (LO), lápiz capacitivo puntero sobre pantalla táctil con medición de la superficie mediante software específico.
Resultados: Se observó una concordancia aceptable entre ambos métodos, coeficiente de correlación intraclase (CCI): 0,87, pero con unos límites de acuerdo excesivamente grandes y un sesgo sistemático consistente en mayores medidas de BSA con MT que con LO. La concordancia entre métodos fue superior en el tronco y extremidades inferiores (CCI > 0,8). La fiabilidad intraobservador fue excelente con ambos métodos (CCI: MT, 0,97; LO: 0,98). La fiabilidad interobservador fue elevada (CCI: MT, 0,91; LO: 0,94), pero el BSA medio difirió significativamente entre observadores. Además, el CCI se redujo drásticamente cuando se consideró la cabeza exclusivamente.
Conclusiones: El presente estudio valida el método ''lápiz óptico'' para la medición de la superficie corporal afectada en pacientes con psoriasis. Muestra una buena concordancia con el método tradicional, presentando menos variabilidad y mayor fiabilidad interobservador.

Introduction

Psoriasis is an inflammatory disease that affects 2%–3% of the adult population in Europe and the United States,[1] with a considerable negative impact on quality of life.[2,3] Currently, there is vigorous debate about whether the severity of psoriasis should be classified in 2 categories (mild and moderate-severe) or 3 categories (mild, moderate, and severe).[4–6] Different scales have been used to provide an objective assessment of disease severity, although none of these fully satisfy the international community.[7,8] The Psoriasis Area and Severity Index (PASI), with a high degree of inter- and intraobserver reliability,[9] is the most widely used and accepted, despite having known limitations.[5] The Dermatology Life Quality Index (DLQI) takes into account the view the patients themselves have of their disease. However, the lack of transcultural equivalence limits the validity of this quality-of-life questionnaire. The third most widely used scale is the Body Surface Area (BSA). The classic (or traditional) method of assessment consists of visual determination of the affected body area, assuming that the area of the palm of the patient accounts for 1% of the body surface area. Therefore, this scale can range from 0% to 100%. The body is subdivided into head and neck (9%), each arm (9%), anterior face (18%) and posterior face

(18%) of the trunk, each leg (18%), and perineum (1%).[10] In agreement with other international groups and expert recommendations,[4,11,12] the Psoriasis Group of the Spanish Academy of Dermatology and Venereology reached a consensus to define severity of psoriasis in function of different criteria, which include PASI, BSA, and DLQI assessed by the same person.[2]

BSA is subjective, tends to overestimate severity, is not particularly useful for some clinical forms (for example, guttate psoriasis), is difficult to assess in patients with residual dyschromia, and the 1% rule is controversial.[12]

The objective of this study was to validate an application for BSA assessment (with an optical pencil [OP]), comparing it with the traditional method ™. The study hypothesis was that the application could represent a more objective assessment of BSA, thus decreasing inter- and intraobserver variability.

Patients and Methods

We conducted a multicenter, prospective, study, which enrolled adult patients with plaque psoriasis who were consecutively seen in different hospitals in Spain and Portugal (Hospital Universitario de la Princesa in Madrid, Hospital General in Valencia, Hospital de Bellvitge in Barcelona, Com-
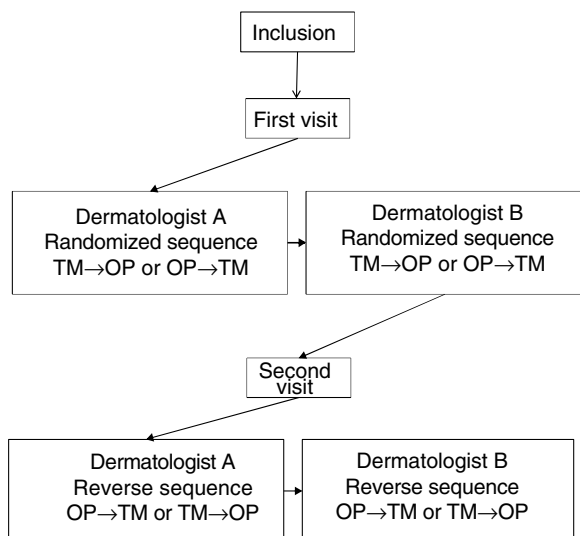
Figure 1 Study design. Abbreviations: OP, optical pencil; TM, traditional method.

plexo Hospitalario Universitario in Santiago de Compostela, Hospital Reina Sofía in Cordoba, all in Spain, and Hospital Santa María in Lisbon, Portugal). The study was approved by the ethics committee of each hospital, and written informed consent was obtained from each patient.

Each patient was assessed at 2 visits, separated by 3 ± 1 days. In the first visit, sociodemographic and clinical variables were recorded. At each visit, 2 dermatologists at the same hospital, blinded to the results of the other dermatologist, performed an assessment of the BSA performed with the TM and the OP application, designed by Serono. Not more than 1 h could elapse between the assessment of the investigators. Fig. 1 shows design of the study.

To use the OP application, the investigator applied a capacitive stylus to a touchscreen showing a representation of a human figure, and shaded the areas affected to obtain a complete planimetry of involved body area (Fig. 2); this information was stored by the application. The order of application of the methods of evaluation of BSA (TM → OP

or OP → TM) was randomized at the first visit and reversed for the second.

Statistical analysis

Measures of BSA were summarized as mean and standard deviation (SD). To study the concordance between methods, the mixed effects intraclass correlation coefficient (ICC) was used for absolute agreement.[13] This same model was used for all analyses except for study of interobserver reliability, for which the ICCs were used as an indicator of absolute agreement. The Bland–Altman graphical method[14] was also used to study the difference between methods, with d the SD of the difference, s the limits of agreement, and $a -1.96s$ and $d +1.96s$ the individual differences between methods. To compare the variability of both methods, the Levene test of equality of variances was used.[15] Finally, analysis of variance (ANOVA) was used with 4 intrasubject factors; method, investigator, visit, and order of application. The main effects and first, second, and third-degree interactions were studied.

A questionnaire with closed dichotomic questions was used to assess the opinions of the observers about the OP. The questionnaire included questions on feasibility, usefulness as a prescribing tool, possible standardization, usefulness in clinical trials, and usefulness in clinical practice.

Statistical significance was set at $P < .05$. Statistical analysis was performed with the SPSS program, version 21.0.

Results

Fifty-six patients were included, 34 men (60.7%) and 22 women (39.3%) with a mean (SD) age of 49.93 (16.76) years. The mean duration of psoriasis from diagnosis was 18.77 (14.28) years. The mean baseline PASI was 11.4 (8.7). Treatments for psoriasis administered were topical therapy in 24 patients (42.9%), phototherapy in 17 patients (30.4%), conventional systemic therapy in 14 patients (25%), and biologic therapy in 16 patients (28.6%).
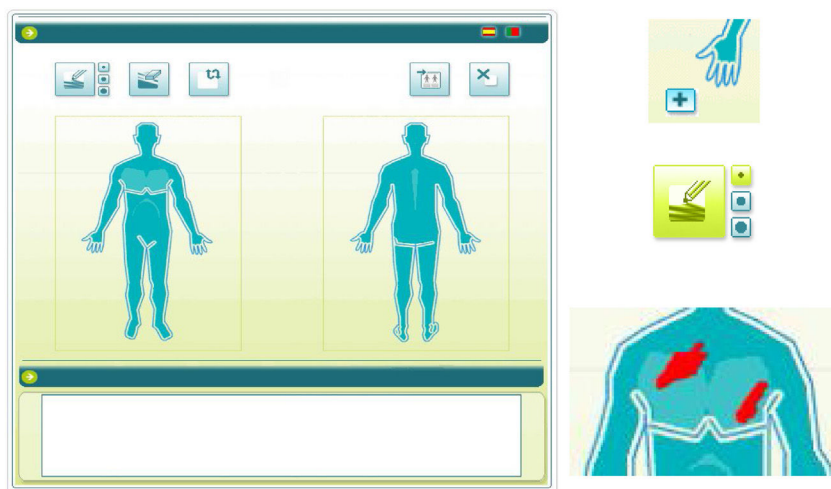
Figure 2 Details of the optical pencil method.
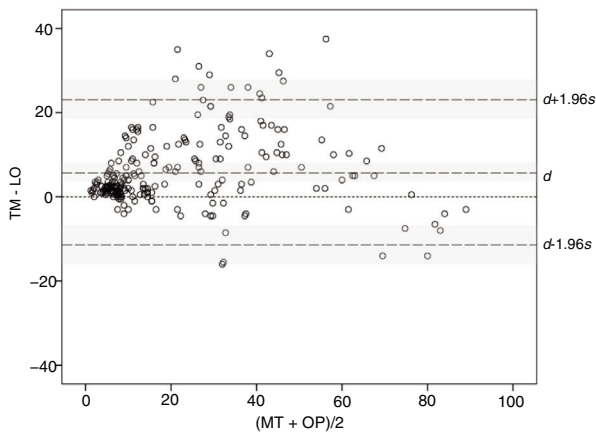
Figure 3 Graphical concordance between the 2 methods. Abbreviations: *d*, mean difference; OP, optical pencil method; *s*, standard deviation of the difference; TM, traditional method.

Concordance between measurement methods

With TM, the mean (SD) BSA obtained in 224 measurements (4 measurements in 56 patients) was 25.63 (20.90), significantly greater than with OP, with a mean value of 19.36 (19.5) ($P < .001$), and a similar trend was observed in the different areas analyzed (Table 1).

The ICC between the 2 methods was 0.87 (95% CI, 0.83–0.90). By body area analyzed, the highest ICCs were found for the trunk and legs, and were lower for the arms and even lower for the head.

The concordance between methods was studied using the Bland–Altman graphical method. The mean difference, 5.85, and the corresponding 95% CI (3.94–8.59) (indicated as *d* and shaded area around this marker in Fig. 3) remained entirely above 0, thereby confirming the presence of a systematic bias, consistent with higher BSA values with TM compared with OP. Moreover, the distribution of the differences suggests that the bias arises with small and median sized BSA values, where BSA with TM is often greater than the value obtained with OP, whereas at very high values, BSA > 70, the tendency is inverted and the values measured with OP are greater than with TM. On the other hand, the limits of agreement, lower limit −11.10 and upper limit 23.55 (indicated with *d* −1.96*s* and *d* +1.96*s* with corresponding 95% CIs [shaded areas] in Fig. 3), are somewhat larger than desired, suggesting a moderate precision in some of the methods.

Intermethod Concordance at Each Visit

The concordance between the 2 methods was studied at each of the visits, using the 2 measurements per patient and method performed at each visit (N = 112). The mean (SD) BSAs for TM and OP were 26.35 (21.66) and 20.14 (17.78), respectively, at the first visit ($P < .001$) and 24.91 (20.21) and 19.43 (19.34), respectively, at the second visit ($P < .001$). At both visits, the mean BSA was greater with TM than with OP.

The ICC between the 2 methods for those patients who had undergone the first assessment of BSA according to the sequence TM → OP (N = 56) was 0.87 (95% CI, 0.82–0.91) and for those patients assessed according to the reverse

sequence (N = 56) the ICC was 0.86 (95% CI, 0.80–0.90). The concordance between the 2 methods at each visit was similar to the overall concordance.

Test−retest Reliability With Each Method

The mean (SD) BSAs for the 2 methods for the 56 patients (N = 112) in the first and second visit for TM were 25.56 (21.01) and 24.91 (20.21), respectively ($P = .128$) while for OP, the values were 19.38 (19.13) and 19.43 (19.34), respectively ($P = .536$). No significant differences were found between mean BSA obtained on Days 1 and 3 for either of the 2 methods. Table 2 shows the ICCs and the corresponding 95% CIs for the first and second visit for each method for the whole body and by each area studied. The ICCs obtained were greater for OP than for TM for the whole body and for 4 of the anatomical areas considered.

Interobserver concordance with each method

The mean (SD) BSAs for the first (A) and second (B) observer for TM (N = 112) were 24.35 (20.48) and 26.92 (21.36), respectively ($P = .012$) while for OP, the values were 18.65 (18.61) and 20.92 (20.42), respectively ($I = .004$), suggesting that the assessments of observer B were greater than those of observer A, regardless of the procedure used. Table 3 shows the ICCs and the corresponding 95% CIs between observers A and B for each method for the whole body and by each area studied. Regardless of the body area under study, the ICCs were greater for OP than for TM. The body area with lowest concordance between dermatologists in both methods was the head.

Analysis of variance

The Levene test for equality of variance showed that the BSA variance was significantly lower when measured with OP than with TM ($P < .001$).

The repeat measure ANOVA only found a main effect for the factor of method ($F_{1. 431} = 8.742$; $P = .003$). None of the remaining principle effects or interactions studied were statistically significant.

Opinion of the investigators on the optical pencil method

Most of the 12 investigators considered that the OP method was an easy-to-use tool (92%), that should be standardized (83%), and that could be used in clinical trials (100%). In contrast, the method received less support for prescribing uses (67%) and use in daily clinical practice (50%).

Discussion

Although there is agreement among dermatologists about the need to use measurement scales in clinical practice, in reality, it is rarely used routinely. The traditional assessment of BSA is based on the subjective assumption of the 1% rule,[13] reducing interobserver reliability, and introducing the possibility of measurement errors.[4,8] Although much progress has

Table 1   Mean Body Surface Area (BSA) of 4 Measurements in 56 Patients (N = 224) With the Traditional Method and Optical Pencil Method for the Entire Body and by Body Areas Analyzed: Intraclass Correlation Coefficient (ICC) Between Mements in 56 patients (N = 224) with the traditional method and optical pencil method for the entire body and by body areas analyzed: intraclass correlation coefficient (ICC) between methods.

Area	Method		Difference Mean (TM)	P	ICC (95% CI)
	Traditional Mean (TM)	Optical Pencil Mean (TM)			
Body as a whole	25.63 (20.90)	19.36 (19.50)	6.27 (8.86)	< .001	0.87 (0.83; 0.90)
Head	1.49 (2.15)	0.79 (1.19)	0.70 (1.49)	< .001	0.57 (0.47; 0.65)
Trunk	8.29 (8.37)	6.97 (7.99)	1.32 (4.18)	.011	0.86 (0.82; 0.89)
Arms	4.42 (4.21)	3.13 (3.63)	1.29 (2.55)	< .001	0.74 (0.68; 0.80)
Legs	9.41 (8.24)	7.94 (7.93)	1.47 (4.23)	.006	0.85 (0.81; 0.88)

Data presented as mean and standard deviation (SD).

Table 2   Intraclass Correlation Coefficient (ICC) Between Visits (N = 112) for Each Method of Assessment of Body Surface Area (BSA), Overall and by Body Areas Analyzed.

Area	Traditional Method ICC (95% CI)	Optical Pencil ICC (95% CI)
Body as a whole	0.975 (0.963; 0.982)	0.985 (0.980; 0.990)
Head	0.789 (0.710; 0.850)	0.870 (0.820; 0.910)
Trunk	0.847 (0.780; 0.890)	0.971 (0.960; 0.980)
Arms	0.872 (0.820; 0.910)	0.952 (0.930; 0.970)
Legs	0.909 (0.870; 0.940)	0.968 (0.960; 0.980)

Table 3   Intraclass Correlation Coefficient (ICC) Between Observers for Each Method of Assessment of Body Surface Area (BSA) and by Body Area Analyzed.

Area	Traditional Method ICC (95% CI)	Optical Pencil ICC (95% CI)
Body as a whole	0.914 (0.880; 0.940)	0.945 (0.920; 0.960)
Head	0.628 (0.50; 0.73)	0.766 (0.67; 0.84)
Trunk	0.895 (0.85; 0.93)	0.932 (0.90; 0.95)
Arms	0.857 (0.80; 0.90)	0.896 (0.85; 0.93)
Legs	0.825 (0.75; 0.88)	0.848 (0.78; 0.89)

been made to automate measurement of PASI,[16–19] few studies have aimed to achieve a more objective measurement of BSA.

The OP method consisted of a system in which the affected surface is shaded on a graphical representation of the body on a touchscreen with a stylus; the application then calculates the percentage of area affected.

Acceptable concordance was observed between TM and OP, although over a wide range of BSA values, TM tends to return lower values than OP, a finding consistent with those of the metaanalysis by Rhodes et al.,[12] demonstrating that BSA based on the 1% rule overestimates the true BSA in adults. On the other hand, the upper limit of agreement of 23.55% estimated using the Bland–Altman graphic method, is high, suggesting a lack of precision in one of the methods,[14] a finding in line with some authors,[15,20] who suggest that TM is not very precise or reproducible. Both methods show greater concordance in large areas such as the trunk and legs compared with smaller areas such as the head and arms, where concordance is lower.

Given that at each visit, the second measurement of each observer was not blinded to the first measurement, a carry-over effect cannot be ruled out. For this reason, the sequence of application of the methods was randomized at the first visit, and reversed for the second. The ICCs obtained in both sequences were similar to one another, suggesting that a carry-over effect was not present.

With regards the test–retest reliability, concordance was measured between the assessments at the first and second visit, showing ICCs greater than 0.95 for the entire body. The test–retest reliability for OP was also excellent for the trunk, legs, and arms, and was slightly lower for the head. These results indicate excellent test–retest reliability.

Ramsay and Lawrence[21] reported BSA measured in a sample of 10 patients ranging from 14% to 33% depending on the observer. More recently, other authors have found modest interobserver reliability for BSA.[8,22,23] In our study, we found significant differences between observers in the mean BSA obtained by the 2 methods. However, the interobserver reliability was good with both methods, with higher ICCs for OP

than for TM, both for the body as a whole and for different areas studied, although the ICC values dropped notably when only the head was considered.

Finally, in terms of the opinions of the participating dermatologists, most considered that the OP method was a tool that was easy to use, could be readily standardized to assess the affected body area, and potentially useful in clinical research. Additionally, OP could be useful for monitoring disease progression in patients, given that readings can be readily stored, and these could be incorporated into electronic medical histories.

An important limitation of this study is that there was no photographic archive to establish a gold standard with which to compare the findings with each of the techniques used. Another limitation is the possible underestimation of BSA with the OP method on transforming a 3D topography onto a 2D projection, in comparison with TM. Likewise, there was no formal measurement of the time used for each of the methods, although the opinion of the investigators was that the time employed for each method was approximately the same or lower for OP method as the investigator was spared the mental calculation required with the TM.

In conclusion, the present study validates the OP method for measurement of BSA in patients with psoriasis. Concordance with TM is not as good as could be expected for interchangeable methods, possibly reflecting inherent limitations of TM. On the other hand, the rest-retest reliability was excellent and the interobserver reliability improved discretely compared with TM. OP is a reliable method that, although potentially applicable in clinical practice, is aimed particularly for use in clinical trials.

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

1. Christophers E. Psoriasis–epidemiology and clinical spectrum. Clin Exp Dermatol. 2001;26:314–20.
2. Dauden E, Puig L, Ferrandiz C, Sanchez-Carazo JL, Hernanz-Hermosa JM. Consensus document on the evaluation and treatment of moderate-to-severe psoriasis: Psoriasis Group of the Spanish Academy of Dermatology and Venereology. J Eur Acad Dermatol Venereol. 2016;30 Suppl. 2:1–18.
3. Finlay AY. Current severe psoriasis and the rule of tens. Br J Dermatol. 2005;152:861–7.
4. Jose RM, Roy DK, Vidyadharan R, Erdmann M. Burns area estimation-an error perpetuated. Burns. 2004;30:481–2.
5. Kimball AB, Jacobson C, Weiss S, Vreeland MG, Wu Y. The psychosocial burden of psoriasis. Am J Clin Dermatol. 2005;6:383–92.
6. Llamas-Velasco M, de la Cueva P, Notario J, Martínez-Pilar L, Martorell A, Moreno-Ramírez D. Moderate psoriasis: a proposed definition. Actas Dermosifiliogr. 2017;108:911–7.
7. Krueger GG, Feldman SR, Camisa C, Duvic M, Elder JT, Gottlieb AB, et al. Two considerations for patients with psoriasis and their clinicians: what defines mild, moderate, and severe psoriasis? What constitutes a clinically significant improvement when treating psoriasis? J Am Acad Dermatol. 2000;43 2 Pt 1:281–5.
8. Puzenat E, Bronsard V, Prey S, Gourraud PA, Aractingi S, Bagot M, et al. What are the best outcome measures for assessing plaque psoriasis severity? A systematic review of the literature. J Eur Acad Dermatol Venereol. 2010;24 Suppl. 2:10–6.
9. Lebwohl M. A clinician's paradigm in the treatment of psoriasis. J Am Acad Dermatol. 2005;53 Suppl. 1:S59–69.
10. Long CC, Finlay AY, Averill RW. The rule of hand: 4 hand areas - 2 FTU - 1g. Arch Dermatol. 1992;128:11–30.
11. Mrowietz U, Kragballe K, Reich K, Spuls P, Griffiths CE, Nast A, et al. Definition of treatment goals for moderate to severe psoriasis: a European consensus. Arch Dermatol Res. 2011;303:1–10.
12. Rhodes J, Clay C, Phillips M. The surface area of the hand and the palm for estimating percentage of total body surface area: results of a meta-analysis. Br J Dermatol. 2013;169:76–84.
13. Koo T, Li M. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155–63.
14. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;327:307–10.
15. Levene H. Robust tests for the equality of variances. In: Olkins J, editor. Contributions to probability and statistics. Palo Alto, CA: Standford University Press.
16. Fink C, Alt C, Uhlmann L, Klose C, Enk A, Haenssle HA. Precision and reproducibility of automated computer-guided Psoriasis Area and Severity Index measurements in comparison with trained physicians. Br J Dermatol. 2019;180:390–6.
17. Fink C, Fuchs T, Enk A, Haenssle HA. Design of an algorithm for automated, computer-guided PASI measurements by digital image analysis. J Med Syst. 2018;42:248.
18. Fink C, Uhlmann L, Klose C, Haenssle HA. Automated, computer-guided PASI measurements by digital image analysis versus conventional physicians' PASI calculations: study protocol for a comparative, single-centre, observational study. BMJ Open. 2018;8:e018461.
19. Kreft S, Kreft M, Resman A, Marko P, Kreft KZ. Computer-aided measurement of psoriatic lesion area in a multicenter clinical trial-comparison to physician's estimations. J Dermatol Sci. 2006;44:21–7.
20. Henseler T, Schmitt-Rau K. A comparison between BSA, PASI, PLASI and SAPASI as measures of disease severity and improvement by therapy in patients with psoriasis. Int J Dermatol. 2008;47:1019–23.
21. Ramsay B, Lawrence CM. Measurement of involved surface area in patients with psoriasis. Br J Dermatol. 1991;124:565–70.
22. Bozek A, Reich A. The reliability of three psoriasis assessment tools: Psoriasis area and severity index, body surface area and physician global assessment. Adv Clin Exp Med. 2017;26:851–6.
23. Yune YM, Park SY, Oh HS, Kim DJ, Yoo DS, Kim IH, et al. Objective assessment of involved surface area in patients with psoriasis. Skin Res Technol. 2003;9:339–42.