

The reliability of three psoriasis assessment tools: Psoriasis area and severity index, body surface area and physician global assessment

Agnieszka Bożek^{A–F}, Adam Reich^{A–F}

Department of Dermatology, Venereology and Allergology, Wrocław Medical University, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of article

Advances in Clinical and Experimental Medicine, ISSN 1899-5276 (print), ISSN 2451-2680 (online)

Adv Clin Exp Med. 2017;26(5):851–856

Address for correspondence

Adam Reich
E-mail: adi_medicalis@go2.pl

Funding sources

None declared

Conflict of interest

None declared

Acknowledgements

The authors would like to thank all the physicians and patients who agreed to voluntarily participate in this study.

Received on December 31, 2016

Revised on February 18, 2017

Accepted on March 27, 2017

Abstract

Background. A wide variety of psoriasis assessment tools have been proposed to evaluate the severity of psoriasis in clinical trials and daily practice. The most frequently used clinical instrument is the psoriasis area and severity index (PASI); however, none of the currently published severity scores used for psoriasis meets all the validation criteria required for an ideal score.

Objectives. The aim of this study was to compare and assess the reliability of 3 commonly used assessment instruments for psoriasis severity: the psoriasis area and severity index (PASI), body surface area (BSA) and physician global assessment (PGA).

Material and methods. On the scoring day, 10 trained dermatologists evaluated 9 adult patients with plaque-type psoriasis using the PASI, BSA and PGA. All the subjects were assessed twice by each physician. Correlations between the assessments were analyzed using the Pearson correlation coefficient. Intra-class correlation coefficient (ICC) was calculated to analyze intra-rater reliability, and the coefficient of variation (CV) was used to assess inter-rater variability.

Results. Significant correlations were observed among the 3 scales in both assessments. In all 3 scales the ICCs were > 0.75, indicating high intra-rater reliability. The highest ICC was for the BSA (0.96) and the lowest one for the PGA (0.87). The CV for the PGA and PASI were 29.3 and 36.9, respectively, indicating moderate inter-rater variability. The CV for the BSA was 57.1, indicating high inter-rater variability.

Conclusions. Comparing the PASI, PGA and BSA, it was shown that the PGA had the highest inter-rater reliability, whereas the BSA had the highest intra-rater reliability. The PASI showed intermediate values in terms of inter- and intra-rater reliability. None of the 3 assessment instruments showed a significant advantage over the other. A reliable assessment of psoriasis severity requires the use of several independent evaluations simultaneously.

Key words: psoriasis, measurement tools, PGA, BSA, PASI

DOI

10.17219/acem/69804

Copyright

© 2017 by Wrocław Medical University
This is an article distributed under the terms of the
Creative Commons Attribution Non-Commercial License
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Psoriasis is a common chronic inflammatory disorder with multiple pathways of pathogenesis that can be associated with metabolic and cardiovascular disease.¹ It is a disease with a high burden and substantial impact on quality of life. Since there are no biomarkers available to assess disease severity, clinical assessment tools are used to measure the disease severity and treatment response in clinical trials and daily practice.² The ideal assessment instrument should consistently represent the true degree of disease severity, minimize inter- and intra-rater variability, be responsive to change, use the entire range of the scale (wide response distribution) and be easy to administer.^{3–5} A wide variety of scoring systems have been proposed to assess the severity of psoriasis. More than 44 different outcome scoring systems were used in 171 randomized clinical trials of psoriasis therapies between 1977 and 2000, as reported in a review by Naldi et al.³ Another systematic review performed in 2010 identified 53 separate clinical assessment instruments.² The psoriasis area and severity index (PASI) is the most commonly used clinical scoring system in research, but it has several disadvantages. The PASI has been criticized for being resource intensive, complex, lacking sensitivity, being low in accuracy and having a non-linear scale. Despite these limitations, the PASI is the most extensively used psoriasis clinical severity score and is considered the reference scoring system against which other assessments tools are compared. Based on systematic reviews, it appears that no single outstanding instrument has been developed for psoriasis, and none of the severity scores used for psoriasis meet all the validation criteria required for an ideal assessment instrument.^{2,6} Therefore, further work is still needed to validate existing instruments and to develop new and better ones.

The aim of this study was to compare and assess the reliability of 3 commonly used measures of psoriasis severity: the PASI, body surface area (BSA) and physician global assessment (PGA).

Material and methods

The subjects

Ten adult patients with plaque-type psoriasis of varying extent were recruited for the study at the Department of Dermatology, Venereology and Allergology, Wrocław Medical University, Poland. All the subjects voluntarily agreed to participate in the study and written informed consent was obtained from all the subjects before any further procedure. Patients with erythrodermic, guttate or pustular psoriasis were excluded. All the participants were instructed not to apply topical treatment prior to the examination until all the assessments were completed. One patient withdrew consent at the

beginning of the experiment and ultimately 9 patients were assessed.

The assessors

Ten practicing dermatologists (7 females and 3 males) were chosen from the Department staff to act as assessors in the study. The physicians' ages ranged from 26 to 36 years, and their duration of dermatology practice ranged from 1 to 11 years. Just before the study, all the participating physicians took part in a training session about using the assessment instruments and recording the data.

The study design

The study was conducted on a single day at Wrocław Medical University's Department of Dermatology, Venereology and Allergology. In total, 9 subjects were evaluated twice by each of 10 physicians.

On the scoring day each patient was randomly allocated an individual examination room, where s/he remained until the study was completed. The patients received numbers; their names were not used. The physicians rotated from one examination room to the next one in a sequential order, which was precisely controlled. At any given time, only one assessor was assigned to one patient. Each physician independently scored the severity of each patient's psoriasis using the PASI, BSA and PGA, using standardized forms. Each observer used one form for each assessment tool per patient. Each form was collected immediately upon completion and stored until the end of the scoring session. The physicians evaluated all the patients with no time limit for the examinations. In general, the time it took each physician to complete all the assessments for all the patients was about 1 h. Directly after the first assessment, each physician was asked to evaluate each patient again, in the same order as previously, without access to the previous scores. The physicians were not allowed to discuss the assessments with one another until all the study forms were completed. Similarly, no communication between the physicians and the patients was allowed during the examinations.

Scoring methods

Psoriasis area and severity index (PASI)

The PASI was developed in 1978 by Fredriksson and Pettersson to assess the effects of retinoids in psoriasis.⁷ The PASI combines assessments of 4 body areas: the head and neck (H), the upper limbs (UL), the trunk (T) and the lower limbs (LL). The percentage of skin affected by psoriasis in each area is given a numerical score (A) representing the proportion involved: 1 (0–9%),

2 (10–29%), 3 (30–49%), 4 (50–69%), 5 (70–89%) or 6 (90–100%). Within each area (H, UL, T, LL) the severity of 3 plaque signs – erythema (E), thickness/induration (I) and desquamation/scaling (D) – is assessed on a 5-point scale: 0 (none), 1 (mild), 2 (moderate), 3 (severe) or 4 (very severe). The final PASI score ranges from 0 to 72, and is calculated using the following formula:

$$\text{PASI} = 0.1 (E_H + I_H + H_H) A_H + 0.2 (E_{UL} + I_{UL} + H_{UL}) A_{UL} + 0.3 (E_T + I_T + H_T) A_T + 0.4 (E_{LL} + I_{LL} + H_{LL}) A_{LL}^{7,8}$$

Physician global assessment (PGA)

The PGA is an average assessment of all psoriatic lesions, based on erythema, scale and induration. It neither quantifies body surface area nor evaluates individual lesion locations.⁹ There are 2 primary PGA forms: a static form, which measures the physician’s impression of the disease at a single point, and a dynamic form, in which the physician assesses global improvement from a baseline.¹⁰ The PGA chosen for this study ranged from 0–5 point, and the following categories were used: 0 = clear, 1 = almost clear, 2 = mild, 3 = moderate, 4 = severe, 5 = very severe.

Body surface area (BSA)

The most commonly used method to estimate the BSA of psoriatic lesions is the “rule of nines”, which was originally developed for estimating the surface area of burns. It is defined as 9% coverage for the head and neck, 9% for each arm, 9% for the anterior and posterior legs, and 9%

for each of 4 trunk quadrants, leaving 1% for the genitalia. The BSA can also be estimated by the number of a patient’s hand areas affected, on the assumption that one “handprint” reflects approximately 1% of BSA.²

Statistical analysis

All the results were analyzed statistically using STATISTICA v. 12.0 software (StatSoft Inc., Tulsa, USA). Mean values and standard deviations were calculated for all the measurements. Possible differences between the first and second assessments were verified with a paired Student’s t-test. Correlations between the parameters studied were calculated with the Pearson correlation test. The intra-class correlation coefficient (ICC) was calculated to analyze intra-rater reliability, with the following ranges: < 0.40 – poor reproducibility; 0.4–0.59 – fair, 0.6–0.74 – good; 0.75–1.0 – very good reproducibility. The coefficient of variation was used to assess inter-rater variability. It was interpreted as follows: 0–20% – slight; 21–40% – moderate; 41–60% – high; > 60% – very high variability. Only p-values less than 0.05 were considered significant.

Results

The mean PASI, BSA and PGA scores given each subject by 10 physicians in 2 assessment sessions are presented in Table 1. There were no significant differences between 2 assessment sessions, except for the 3rd patient’s PASI score (p = 0.006).

Table 1. Mean scores for the psoriasis area and severity index (PASI), the body surface area (BSA) and the physician global assessment (PGA) given by 10 physicians in 2 assessments (mean ± standard deviation)

| Patients | PASI | | p-value | BSA | | p-value | PGA | | p-value |
|-----------|--------------|--------------|---------|--------------|--------------|---------|--------------|--------------|---------|
| | assessment 1 | assessment 2 | | assessment 1 | assessment 2 | | assessment 1 | assessment 2 | |
| Patient 1 | 7.5 ± 2.0 | 7.4 ± 2.8 | 0.76 | 8.1 ± 2.9 | 8.0 ± 2.9 | 0.89 | 3.1 ± 0.9 | 3.2 ± 0.9 | 0.59 |
| Patient 2 | 2.3 ± 1.0 | 2.6 ± 1.2 | 0.19 | 5.0 ± 4.7 | 4.2 ± 3.8 | 0.4 | 1.2 ± 0.4 | 1.2 ± 0.4 | 1.0 |
| Patient 3 | 15.0 ± 5.2 | 18.6 ± 6.2 | 0.006 | 40.9 ± 14.3 | 39.1 ± 13.7 | 0.54 | 3.2 ± 0.6 | 3.5 ± 0.5 | 0.19 |
| Patient 4 | 7.4 ± 3.8 | 6.7 ± 3.7 | 0.51 | 20.2 ± 11.4 | 20.1 ± 16.0 | 0.98 | 2.2 ± 1.0 | 2.6 ± 0.5 | 0.35 |
| Patient 5 | 3.2 ± 1.2 | 3.2 ± 1.0 | 0.87 | 3.0 ± 3.5 | 2.8 ± 3.4 | 0.58 | 1.8 ± 0.9 | 1.9 ± 0.9 | 0.34 |
| Patient 6 | 11.8 ± 5.6 | 13.0 ± 6.7 | 0.53 | 39.4 ± 23.5 | 41.1 ± 20.4 | 0.63 | 3.0 ± 1.1 | 3.1 ± 0.9 | 0.68 |
| Patient 7 | 14.2 ± 5.1 | 16.5 ± 6.5 | 0.33 | 28.5 ± 16.5 | 34.9 ± 25.8 | 0.5 | 3.5 ± 0.5 | 3.4 ± 0.7 | 0.59 |
| Patient 8 | 29.6 ± 6.0 | 27.0 ± 11.1 | 0.44 | 90.2 ± 13.4 | 84.8 ± 24.3 | 0.62 | 3.8 ± 0.6 | 4.0 ± 0.7 | 0.17 |
| Patient 9 | 11.4 ± 3.8 | 12.9 ± 4.8 | 0.4 | 23.2 ± 9.3 | 20.9 ± 8.1 | 0.09 | 3.4 ± 0.7 | 3.5 ± 0.5 | 0.59 |

Table 2. Pearson correlation coefficients (*r*) between the psoriasis area and severity index (PASI), the body surface area (BSA) and the physician global assessment (PGA) obtained by 10 physicians in 2 assessments ($p < 0.05$ for all *r* values)

| Assessment | Scale | Assessment 1 | | | Assessment 2 | | |
|--------------|-------|--------------|------|------|--------------|------|------|
| | | PASI | BSA | PGA | PASI | BSA | PGA |
| assessment 1 | PASI | – | 0.84 | 0.61 | 0.85 | 0.77 | 0.64 |
| | BSA | – | – | 0.55 | 0.68 | 0.87 | 0.56 |
| | PGA | – | – | – | 0.53 | 0.53 | 0.85 |
| assessment 2 | PASI | – | – | – | – | 0.78 | 0.66 |
| | BSA | – | – | – | – | – | 0.61 |

Significant correlations were observed among the scales in both assessments (Table 2). The Pearson correlation coefficient between the PASI and BSA for all 10 physicians indicates that the evaluations were very strongly correlated in the 1st assessment (0.84) and strongly correlated in the 2nd assessment (0.78). The correlation between the PASI and PGA was strong in both the 1st and 2nd assessments (0.61 vs 0.66). The correlation between the PGA and BSA was moderate in the 1st assessment (0.55) and strong in the 2nd assessment (0.61).

Table 3 summarizes the intra-rater reliability of the PASI, BSA and PGA for all the physicians. In all 3 scales the ICCs were > 0.75 , indicating very good reliability. The highest ICC was for the BSA (0.96) and the lowest

one was for the PGA (0.87). The ICC for all components of the PASI were also > 0.75 , indicating very good reliability, except for the ICC for scaling (0.72), indicating good reliability (Table 3). Among the components of the PASI, the highest ICC was observed for the area score (0.97).

Table 4 shows the inter-rater variability of the PASI, BSA and PGA for all the physicians. The coefficients of variation (CVs) for the PGA and PASI were 29.3 and 36.9, respectively, indicating moderate variability. The CV for the BSA was 57.1, indicating high variability. The CVs for the separate components of the PASI showed the highest variability for the head and neck (117.8) and the lowest 1 for the area score (26.8) (Table 4).

Table 3. Intra-class correlation coefficients (ICCs) for the psoriasis area and severity index (PASI), each of the PASI components, the body surface area (BSA) and the physician global assessment (PGA)

| Scale | Intra-class correlation coefficient (ICC) | Intra-rater reliability (interpretation) | |
|-------|---|--|-----------|
| PASI | total scoring | 0.91 ± 0.06 | very good |
| | erythema | 0.81 ± 0.16 | very good |
| | induration | 0.77 ± 0.14 | very good |
| | scaling | 0.72 ± 0.21 | good |
| | area score | 0.97 ± 0.02 | very good |
| | head and neck | 0.85 ± 0.13 | very good |
| | trunk | 0.86 ± 0.18 | very good |
| | upper limbs | 0.84 ± 0.11 | very good |
| | lower limbs | 0.89 ± 0.06 | very good |
| | BSA | 0.96 ± 0.05 | very good |
| PGA | 0.87 ± 0.07 | very good | |

Table 4. Coefficients of variation (CVs) for the psoriasis area and severity index (PASI), each of the PASI components, the body surface area (BSA) and the physician global assessment (PGA)

| Scale | Coefficient of variation % | Inter-rater variability (interpretation) | |
|-------|----------------------------|--|-----------|
| PASI | total scoring | 36.9 ± 9.8 | moderate |
| | erythema | 33.2 ± 7.6 | moderate |
| | induration | 35.9 ± 13.9 | moderate |
| | scaling | 59.6 ± 10.3 | high |
| | area score | 26.8 ± 10.3 | moderate |
| | head and neck | 117.8 ± 115.4 | very high |
| | trunk | 59.7 ± 43.0 | high |
| | upper limbs | 49.2 ± 16.7 | high |
| | lower limbs | 42.9 ± 9.7 | high |
| | BSA | 57.1 ± 31.8 | high |
| PGA | 29.3 ± 12.4 | moderate | |

Discussion

The PASI, BSA and PGA are the most frequently used psoriasis assessment tools. While these scales are commonly used to assess psoriasis severity in both clinical research and clinical practice, they have not been fully evaluated in terms of their validity and reliability, and no consensus exists on the best way to monitor psoriasis severity.^{6,8} Despite its widespread use, the PASI has never been formally validated. Its grading has never been properly defined or standardized.¹¹ Therefore, this study was undertaken with the aim of assessing the reliability of the PASI, BSA and PGA.

In the present study a significant correlation was found between all 3 of the psoriasis assessment tools tested. The correlation between the PASI and BSA was higher than between the other scales, whereas the lowest correlation was noted between the BSA and PGA. The lack of assessment of the area affected as a component of the PGA, in contrast to the PASI, is probably responsible for the lower correlation between the PGA and other scales. However, Langley and Ellis reported that the sum of the area scores using PASI was more highly correlated with PGA ($r = 0.8$) than the summed scores for erythema, induration or desquamation ($r = 0.3$ to 0.6), which suggests that physicians do incorporate the extent of involvement when using the PGA, even though it is not stated in the PGA.¹³ However, it must be mentioned that many different versions of the PGA are described in clinical trials, and this variation has the potential to produce non-standardized results. The scales used may range from 4 points to 10 points and the definitions used for the specific points on the scale vary slightly.⁹

The present study demonstrated very good intra-rater reliability for all 3 scales, which can be ranked in the following order from the highest to the lowest In terms of intra-rater reliability: BSA > PASI > PGA. Intra-rater reliability for the BSA has also previously been described as very good.^{14,15} Two studies conducted by Berth-Jones et al. have demonstrated substantial intra-rater reliability for the PASI and PGA, and higher reproducibility for the PASI than for the PGA.^{8,12} However, the data presented by Langley and Ellis revealed over 12 times higher intra-rater variation for the PASI than for the PGA.¹³ The difference between these studies might be due to the statistical methodology.

In the present study, all of the PASI components had very good reproducibility, except for scaling, which had good reproducibility. As compared to the BSA, the PASI area score demonstrated the highest intra-rater reliability among all the PASI components. In terms of intra-rater reliability, the psoriatic plaque features are ranked from the highest to the lowest as follows: erythema > induration > scaling. The body areas are ranked in the following order: lower limbs > trunk > head and neck > upper limbs.

In the present study, the BSA demonstrated high inter-rater variability. A systemic review performed by Puzenat et al. also showed that inter-rater variation was unacceptably high for the BSA.⁶ For this reason, the BSA was not recommended for assessment of the clinical severity of psoriasis. The BSA could be calculated either by the “rule of nines” method or by the number of patient’s hand areas affected. The findings of a meta-analysis performed by Rhodes et al. demonstrated that the widely accepted value of hand surface area as 1% of total BSA is inaccurate.¹⁶ If hand surface area is to be used for adults, the estimate should be 0.9% for men and 0.85% for women. However, these values will be influenced by BMI and ethnic origin. This may explain why the BSA is often overestimated and why this assessment tool shows high inter-rater variability.

In the present study, the PASI and PGA achieved moderate inter-rater variability. In 2 previous studies inter-rater reliability was substantial for the PASI and moderate for the PGA.^{8,12} In contrast to the present study, the PASI was suggested as a more reliable assessment instrument than the PGA. However, in another study the PASI showed more variability than the PGA.¹³

The data from the present study suggests that the head and neck area is the most difficult component of the PASI to be assessed properly, as it demonstrated very high inter-rater variability. This fact was not surprising, because scalp assessment is time-consuming and difficult due to the presence of hair. All other body areas demonstrated high inter-rater variability. It was surprising that, in contrast to the BSA, the body area score showed the lowest intra-rater variability among all the PASI components. The different methods of evaluating the affected area used in the PASI and BSA may be an explanation for this discrepancy. In the BSA the area is normally expressed as a percentage, whereas in the PASI it is given a value on a scale of 1–6 for 4 separate body areas.

Assessments of erythema and induration in the PASI revealed moderate inter-rater variability, whereas scaling showed high variability. There is no agreement on how scaling should be scored. Although the original publication⁷ is used as the standard reference source, it is not desquamation (the shedding of skin scales) but scale thickness that is generally scored.¹¹ Scaling is a very unstable sign; applications of topical treatment and emollients have an influence on it. This was eliminated in the present study, because patients were asked not to apply any topical treatment prior to the examination. Patients rate degree of scaling as a strong indicator of disease severity, so appropriate evaluations of scaling by physicians is important in assessing the severity of psoriasis.²

The relatively high inter-rater variation of the components of the PASI suggests that this scoring method is not suitable for follow-up of isolated signs of intensity and body areas.

The main limitation of this study is its small sample size (9 patients). However, in this study each assessor had to re-

cord over 50 measurements for each patient. Thus, a larger number of patients in a relatively short time period would fatigue the assessors and yield inaccurate results.

In general, comparing the PASI, PGA and BSA showed that the PGA had the lowest (i.e. moderate) inter-rater variability among 3 scales, whereas the BSA had the highest intra-rater reliability. The PASI showed intermediate values in terms of inter- and intra-rater reliability. The PGA is a quick and easy method to perform, making it a suitable tool in everyday practice, where there are different physicians performing each assessment. The BSA is the most reproducible method, which means it should be chosen in situations where repeated assessment are performed by the same physician. In conclusion, none of the 3 scoring methods showed a significant advantage over the others. A reliable assessment of psoriasis severity requires the use of several independent evaluations simultaneously.

References

1. Kupetsky EA, Keller M. Psoriasis vulgaris: An evidence-based guide for primary care. *J Am Board Fam Med.* 2013;26:787–801.
2. Spuls PI, Lecluse LL, Poulsen ML, Bos JD, Stern RS, Nijsten T. How good are clinical severity and outcome measures for psoriasis?: Quantitative evaluation in a systematic review. *J Invest Dermatol.* 2010;130:933–943.
3. Naldi L, Svensson A, Diepgen T, et al. Randomized clinical trials for psoriasis 1977–2000: The EDEN survey. *J Invest Dermatol.* 2003;120:738–741.
4. Chularojanamontri L, Griffiths CE, Chalmers RJ. Responsiveness to change and interpretability of the simplified psoriasis index. *J Invest Dermatol.* 2014;134:351–358.
5. Naldi L. Scoring and monitoring the severity of psoriasis. What is the preferred method? What is the ideal method? Is PASI passé? Facts and controversies. *Clin Dermatol.* 2010;28:67–72.
6. Puzenat E, Bronsard V, Prey S. 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):10–16.
7. Fredriksson T, Pettersson U. Oral treatment of pustulosis palmo-plantaris with a new retinoid, Ro 10-9359. *Dermatologica.* 1979;158:60–64.
8. Berth-Jones J, Grotzinger K, Rainville C, et al. A study examining inter- and intrarater reliability of three scales for measuring severity of psoriasis: Psoriasis area and severity index, physician global assessment and lattice system physician global assessment. *Br J Dermatol.* 2006;155:707–713.
9. Robinson A, Kardos M, Kimball AB. Physician global assessment (PGA) and psoriasis area and severity index (PASI): Why do both? A systematic analysis of randomized controlled trials of biologic agents for moderate to severe plaque psoriasis. *J Am Acad Dermatol.* 2012;66:369–375.
10. Feldman S, Krueger G. Psoriasis assessment tools in clinical trials. *Ann Rheum Dis.* 2005;64(Suppl 2):65–68.
11. Chalmers RJ. Assessing psoriasis severity and outcomes for clinical trials and routine clinical practice. *Dermatol Clin.* 2015;33:57–71.
12. Berth-Jones J, Thompson J, Papp K. Copenhagen Psoriasis Working Group. A study examining inter-rater and intrarater reliability of a novel instrument for assessment of psoriasis: The Copenhagen psoriasis severity index. *Br J Dermatol.* 2008;159:407–412.
13. Langley RG, Ellis CN. Evaluating psoriasis with psoriasis area and severity index, psoriasis global assessment, and lattice system physician global assessment. *J Am Acad Dermatol.* 2004;51:563–569.
14. Ramsay B, Lawrence CM. Measurement of involved surface area in patients with psoriasis. *Br J Dermatol.* 1991;124:565–570.
15. Yune YM, Park SY, Oh HS, et al. Objective assessment of involved surface area in patients with psoriasis. *Skin Res Technol.* 2003;9:339–342.
16. 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.