

Inter-rater concordance study of the PASI (Psoriasis Area and Severity Index) *

Importância da variação do PASI realizado por diversos observadores

Joana Ribeiro Costa de Faria ¹
Luiz Miguel Zabaleta Jimenez ³
João Carlos Regazzi Avelleira ⁵

Aline Rezende Aarão ²
Oscar Hernández Silva ⁴

Abstract: BACKGROUND: Developed in 1978, the PASI (Psoriasis Area and Severity Index) is a method to rate psoriasis severity. Since then, it has been criticized for being extremely long, complex and for yielding significantly divergent interpretations. OBJECTIVE: To determine the PASI inter-rater reliability among three independent evaluators examining the same patients.

METHODS: Cross-sectional study conducted in 2007 at a psoriasis treatment center, where 20 patients were selected. The patients were evaluated by three postgraduate students of Dermatology who independently determined the PASI of each patient. The inter-rater reliability coefficient was determined by employing intraclass correlation coefficients (ICC) and *Bland & Altman plots*.

RESULTS: An ICC of 0.729 (IC 95%: 0.440 – 0.882) showed a significant degree of concordance among evaluators. By graphic means, the *Bland & Altman* plot method presents concordance gap ranges. As shown in this study, narrow gap ranges indicate a good concordance degree. It was also observed that for high PASI values, wider concordance gap ranges predominate.

CONCLUSION: Based on this analysis, we may conclude that the PASI is a reliable indicator of psoriasis severity because it shows significant concordance when independent evaluations are performed.

Keywords: Cross-sectional studies; Psoriasis; Reproducibility of results

Resumo: FUNDAMENTO: O índice da área e gravidade da psoríase foi método desenvolvido em 1978, para quantificar a gravidade da psoríase. Desde então, têm recebido algumas críticas, por ser longo, complexo e apresentar divergências na sua interpretação.

OBJETIVO: Demonstrar o grau de concordância do índice da área e gravidade da psoríase aplicado em pacientes de psoríase por três avaliadores isoladamente.

MATERIAL E MÉTODO: Estudo observacional, transversal, realizado no ambulatório de psoríase, no ano de 2007, onde foram selecionados 20 pacientes. Os pacientes foram avaliados por três pós-graduandos de dermatologia, que determinaram o índice da área e gravidade da psoríase, isoladamente. A fidedignidade interobservador foi avaliada pelo coeficiente de correlação intraclassas (ICC) e os gráficos de *Bland & Altman*.

RESULTADOS: O ICC de 0,729 (IC 95%: 0,440 - 0,882) demonstrou que existe boa concordância entre os observadores. Pela análise dos gráficos de *Bland e Altman*, no máximo, duas diferenças estão fora dos limites de concordância, expressando boa concordância. Observou-se também que existe comportamento sistemático nas diferenças do índice da área e gravidade da psoríase em relação a sua média, segundo o qual mostrou para valores altos do índice da área e gravidade da psoríase, ou seja, um predomínio de diferenças maiores.

CONCLUSÃO: Notamos, através deste estudo, que o índice da área e gravidade da psoríase apresenta bons critérios de confiabilidade quando houve a avaliação da gravidade da psoríase, mesmo que realizado por diferentes avaliadores.

Palavras-chave: Estudos transversais; Psoríase, Reprodutibilidade dos testes

Received on 01.09.2008.

Approved by the Advisory Board and accepted for publication on 18.12.08.

* Work conducted at Prof. Rubem David Azulay Institute of Dermatology, Santa Casa de Misericórdia do Rio de Janeiro - Rio de Janeiro (RJ), Brazil.

Conflict of interest: None / *Conflito de interesse: Nenhum*

Financial funding: None / *Suporte financeiro: Nenhum*

¹ MD, Postgraduate student of Dermatology, Prof. Rubem David Azulay Institute of Dermatology, Santa Casa de Misericórdia do Rio de Janeiro – Rio de Janeiro (RJ), Brazil.

² MD, Postgraduate student of Dermatology, Prof. Rubem David Azulay Institute of Dermatology, Santa Casa de Misericórdia do Rio de Janeiro – Rio de Janeiro (RJ), Brazil.

³ MD, Postgraduate student of Dermatology, Prof. Rubem David Azulay Institute of Dermatology, Santa Casa de Misericórdia do Rio de Janeiro – Rio de Janeiro (RJ), Brazil.

⁴ MD, Postgraduate student of Dermatology, Prof. Rubem David Azulay Institute of Dermatology, Santa Casa de Misericórdia do Rio de Janeiro – Rio de Janeiro (RJ), Brazil.

⁵ Ph.D. in Medicine; Associate Professor at Prof. Rubem David Azulay Institute of Dermatology, Santa Casa de Misericórdia do Rio de Janeiro - Rio de Janeiro (RJ), Brazil.

INTRODUCTION

Psoriasis is a chronic inflammatory disease that involves the skin and articulations. It is universally distributed and affects from 1 to 3% of the population.¹ Most cases are benign and recurrent, but the disease can be physically or socially debilitating.²

The Psoriasis Area Severity Index (**PASI**) is a tool developed in 1978 by Fredricksson and Pettersson to evaluate the severity of psoriasis. This index analyzes the four regions of the body (head, trunk, upper and lower limbs) in relation to erythema, induration (thickness), desquamation (scaling) of the plaques, and body surface area (BSA) affected.³ Scores from 0 to 4 are attributed to erythema, induration and scaling, based on their intensity (0-absent, 1-mild, 2-moderate, 3-intense, 4-severe). The BSA affected is expressed in percentage (1 to 100%), with corresponding scores that range from 1 to 6. Therefore, the score of 1 means less than 10% of area involved; 2 means from 10-29%; 3 from 30-49%; 4 from 50-69%; 5 from 70-89%; 6 from 90-100%.

The **PASI** is the most used tool in clinical trials, but it has been criticized for its non-linear scale, complexity, and length.²

The **PASI** scores that classify psoriasis as severe vary constantly. To some authors, the severity of the **PASI** has been defined as ≥ 18 , while others suggest ≥ 12 . Recently, the clinical guidelines of the British Association of Dermatologists have indicated that eligible patients for the use of biological agents should have a **PASI** greater than 10 and a **DLQI** (Dermatologic Life Quality Index) greater than 15. Other authors argue that a **DLQI** > 10 suggests moderate/severe disease in the "rule of tens".⁴

The result of the **PASI** may vary from 0 to 72, although most patients fall between 0-15³ and very few individuals are classified above this value range.

OBJECTIVE

To investigate the inter-rater reliability of the **PASI** used in clinical trials for assessing disease severity; three researchers with similar knowledge and experience separately evaluated the patients.

MATERIAL AND METHODS

This was an observational, clinical-epidemiological, transversal study conducted at a psoriasis ambulatory, from August to October 2007. Twenty patients were randomly selected, aged from 15 to 70 years, with mild, moderate, and severe psoriasis.

Three first-year postgraduate students of Dermatology evaluated the patients and determined their **PASI** separately. Next, results were obtained and the indexes attributed to each patient were compared, according to each rater.

The inter-rater reliability of the **PASI** was evaluated based on two different aspects. First, by the intraclass correlation coefficient (ICC),⁵ which allowed the assessment of the consistency or reproducibility of the measurements; then, by Bland & Altman plots,⁵ which made possible a graphic analysis of the differences observed among the raters. The level of significance adopted was 0.05. The statistical analysis was done with the software package *Statistical Product and Service Solutions* (version 13.0 for windows, SPSS, Inc.).

RESULTS

First, we outlined the general profile of the case study, showing the mean value of the **PASI** for the 20 patients studied, according to each rater (Table 1).

Analysis of inter-rater reliability was done by intraclass correlation coefficient (ICC), which verified if there is significant agreement among the three raters in the **PASI**. Table 2 presents the ICC, the confidence interval (CI 95%) and its respective significance level (p value) for each pair of raters.

We should note that the closer to (1) the ICC is, the stronger (or more perfect) the agreement among the raters is; the closer to (0), the greater the discrepancy.

Therefore:⁵

ICC $\leq 0.20 \Rightarrow$ no agreement

$0.20 < \text{ICC} \leq 0.40 \Rightarrow$ weak agreement

$0.40 < \text{ICC} \leq 0.60 \Rightarrow$ moderate agreement

$0.60 < \text{ICC} \leq 0.80 \Rightarrow$ good agreement

ICC $> 0.80 \Rightarrow$ excellent agreement

Our results show that there is good agreement among the three raters; raters 2 and 3 showed excellent agreement, with an ICC of 0.817 (CI 95%: 0.601-0.923).

Agreement among the raters verified by Bland & Altman plots assessed the dispersion of the differences among the raters against its respective mean value. The amplitude of the agreement intervals indicated the quality of the concordance among the raters. In other words, relatively wide intervals express weak agreement and narrow intervals indicate good agreement among them. Moreover, we can identify a systematic behavior of the differences that may suggest an under or overestimation in relation to the second observer. The limits (upper and lower) correspond to the limits of agreement (Table 3, Graphs 1-3).

By perceptively analyzing the graphs, we may note:

The mean of the differences among the three raters is very close to zero (less than two points), suggesting that the raters are replicating one another.

Only two differences are beyond the limits of agreement, expressing good quality of concordance,

TABLE 1: General descriptive analysis of the PASI

Rater	n	Mean	SD	Median	Minimum	Maximum
1	20	14,4	14,5	9,5	0,8	58,8
2	20	15,3	13,5	11,5	1,2	44,4
3	20	15,8	14,8	10,2	0,6	43,5

SD: Standard deviation

even though the limits are not very narrow.

There is a **systematic behavior of the differences of the PASI** in relation to its mean, which showed a predominance of greater differences for higher values of the PASI and of smaller ones for lower values. In other words, the raters showed **greater** differences for more severe patients. This is a negative aspect because the expected behavior is a random distribution of the differences around the general mean.

The best agreement, based on Bland & Altman plots, was between raters 2 and 3, consistent with the ICC analysis.

DISCUSSION

Despite the importance of evaluating the evolution of psoriasis, there is no consensus about the more appropriate indexes and, unfortunately, there are very few randomized clinical trials that prove the reliability of the methods employed. ⁶

The PASI is the most commonly used tool; however, it is not fully satisfactory in terms of validity and reliability. ⁶

“Validity” may be defined as the capacity of the scale to assess what it is supposed to measure. “Reliability” is the extension to which the same results are obtained under the same conditions. Important reliability measurements for severity scales include inter and intrarater reliability. ⁴

Inter-rater reliability represents the level of agreement among different raters applying the scale to assess the same patient. ⁴

Intra-rater reliability is the degree of agreement

among multiple repetitions of a test performed by a single rater. ⁴

In this study we noticed that there is agreement among the raters regarding the application of the method, but greater differences predominate for higher PASI scores. Nonetheless, since the PASI scores that indicate severity are between 10 and 20, the decision about the need for a more efficient therapy for severe patients would not be affected.

It is possible that studies that test the therapeutic efficiency of new drugs to treat patients with very high PASI scores show differences when conducted by different raters.

Even though the PASI is one of the main tools to evaluate the severity of psoriasis, the index has its limitations, especially when used as a single parameter to orient the medical decision.

The PASI does not efficiently translate a significant reduction of the erythema, scaling and infiltration of lesions when there is no modification of the body surface area (BSA) affected. Other symptoms and signs of the severity of psoriasis should be considered in the evaluation: pruritus, pain or burning on the site of lesions, duration of remission, and presence of articular disease.

Last, it is important to mention that the PASI identifies only the physical characteristics of the lesions and the area affected by psoriasis, but it does not take into account the perception that the patients have of the disease and the impact it has on their quality of life. ⁷

TABLE 2: Intraclass correlation coefficient

Comparison	n	ICC	CI 95%	p value
Obs 1 x Obs 2	20	0,729	0,440 - 0,882	0,00007
Obs 1 x Obs 3	20	0,753	0,481 - 0,894	0,00003
Obs 2 x Obs 3	20	0,817	0,601 - 0,923	< 0,00001

ICC: Intraclass coefficient

CI 95%: Confidence interval of 95% for the ICC.

TABLE 3: Mean and limits of agreement

Rater	n	Mean	SD	LL 95%	UL 95%
1 x 2	20	-0,895	10,52	-21,52	19,73
1 x 3	20	-1,365	10,45	-21,84	19,11
2 x 3	20	-0,470	8,73	-17,59	16,65

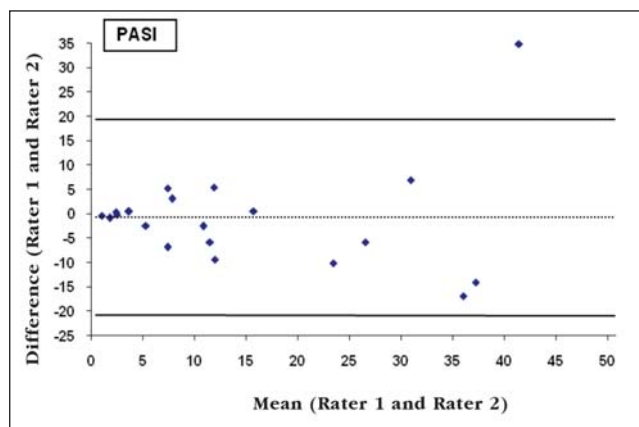
Mean: mean of the differences between raters

SD: Standard deviation of the difference

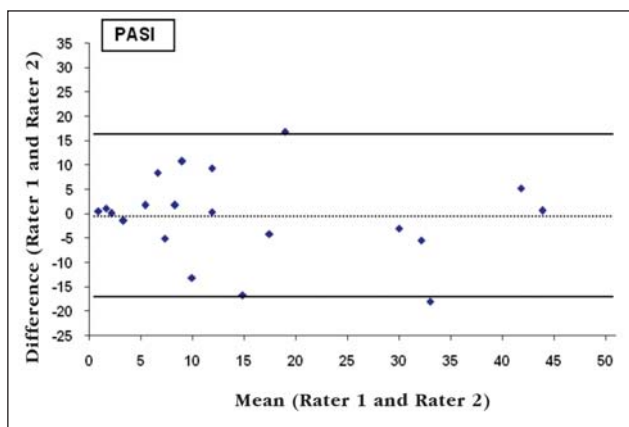
LL 95%: Lower limit of 95% for the difference.

UL 95%: Upper limit of 95% for the difference.

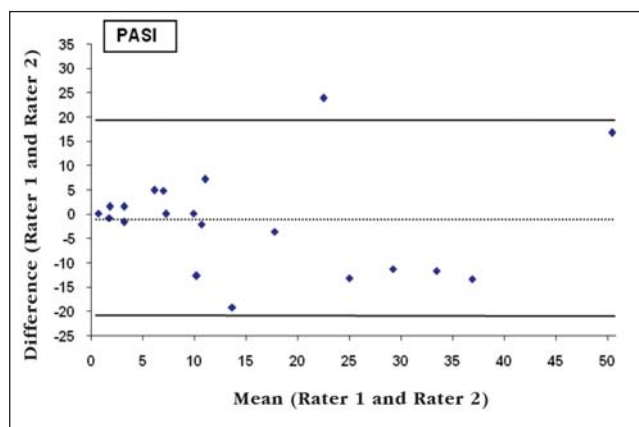
GRAPH 1: Bland-Altman plot for the PASI (rater 1 x rater 2)



GRAPH 3: Bland-Altman plot for the PASI (rater 2 x rater 3)



GRAPH 2: Bland-Altman plot for the PASI (rater 1 x rater 3)



CONCLUSION

The PASI shows good reliability criteria in the evaluation of the severity of psoriasis, despite the fact that it may present some differences of interpretation for patients with higher PASI scores.

We emphasize the need for other variables to be assessed, such as the presence of psoriatic arthritis, the psychosocial aspect of the disease and its impact on the individual's quality of life. These variables, together with the PASI, would influence the therapeutic decision of the assistant physician. □

REFERENCES

1. Azulay RD, Azulay DR. Dermatologia. 4 ed. Rio de Janeiro: Guanabara Koogan; 2006.
2. Carlin CS, Feldman SR, Krueger JG, Menter A, Krueger GG. A 50% reduction in the Psoriasis Area and Severity Index (PASI 50) is a clinically significant endpoint in the assessment of psoriasis. J Am Acad Dermatol. 2004; 50:859-66.
3. Jacobson CC, Kimball AB. Rethinking the psoriasis area and severity index: the impact of area should be increased. Br J Dermatol. 2004; 151:381-7.
4. Berth-Jones J, Grotzinger K, Rainville C, Pham B, Huang J, Daly S, et al. A study examining inter- and intrarater reliability of three scales for measuring severity of psoriasis: Psoriasis Area and Severity Index, Physician's Global Assessment and Lattice System Physician's Global Assessment. Br J Dermatol. 2006; 155:707-13.
5. Bartko JJ, Carpenter WT Jr. On the methods and theory of reliability. J Nerv Ment Dis. 1976;163:307-17.
6. Langley RG, Ellis CN. Evaluating psoriasis with Psoriasis Area and Severity Index, Psoriasis Global Assessment, and Lattice System Physician's Global Assessment. J Am Acad Dermatol. 2004; 51:563-9.
7. Grob JJ. Why are quality of life instruments not recognized as reference measures in therapeutic trials of chronic skin disorders? J Invest Dermatol. 2007; 127:2299-301.

ENDEREÇO PARA CORRESPONDÊNCIA / MAILING ADDRESS:

Joana Ribeiro Costa de Faria
 Rua Belfort Roxo, 407 - aptº 601 - Copacabana -
 22020-010 Rio de Janeiro - RJ, Brazil
 Phone: +55 021 2275 7031
 Mobile: +55 021 8159 5206
 Fax: +55 021 2220 1928(IDPRDA)
 E-mail: joanarcf@gmail.com

How to cite this article/Como citar este artigo: Faria JRC, Aarão AR, Jimenez LMZ, Silva OH, Avelleira JCR. Inter-rater concordance study of the PASI (Psoriasis Area and Severity Index). An Bras Dermatol. 2010;85(5):625-9.