Large International Validation of ABSIS and PDAI Pemphigus Severity Scores

Vivien Hébert, Claire Boulard, Estelle Houivet, Sophie Lehembre, Luca Borradori, Rocco Della, Claudio Feliciani, Luca Fania, Giovanna Zambruno, Diana Camaioni, et al.

Large International Validation of ABSIS and PDAI Pemphigus Severity Scores

Vivien Hébert[1], Claire Boulard[1], Estelle Houivet[2], Sophie Duvert Lehembre[3], Luca Borradori[4], Rocco Della Torre[4], Claudio Feliciani[5], Luca Fania[6], Giovanna Zambruno[7], Diana B. Camaioni[7], Biago Didona[7], Branka Marinovic[8], Enno Schmidt[9], Nina Schumacher[9], Christian Hünefeld[10], Stefan Schanz[10], Johannes Steffen Kern[11], Silke Hofmann[11], Anne Charlotte Bouyeure[2], Catherine Picard-Dahan[12], Catherine Prost-Squarcioni[13], Frederic Caux[13], Marina Alexandre[13], Saskia Ingen-Housz-Oro[14], Martine Bagot[15], Emmanuelle Tancrede-Bohin[15], Jean David Bouaziz[15], Nathalie Franck[16], Pierre Vabres[17], Bruno Labeille[18], Marie Aleth Richard[19], Emmanuel Delaporte[20], Alain Dupuy[21], Michel D'Incan[22], Gaelle Quereux[23], François Skowro[24], Carle Paul[25], Cristina Bulai Livideanu[25], Marie Beylot-Barry[26], Marie Sylvie Doutre[26], Martine Avenel-Audran[27], Christophe Bedane[28], Philippe Bernard[29], Laurent Machet[30], Hervé Maillard[31], Denis Jullien[32], Sebastien Debarbieux[32], Bruno Sassolas[33], Laurent Misery[33], Claire Abasq[33], Olivier Dereure[34], Philippe Lagoutte[2], Vincent Ferranti[1], Victoria P. Werth[35], Dedee F. Murrell[36], Michael Hertl[37], Jacques Benichou[2] and Pascal Joly[1], the French Study Group on Autoimmune Bullous Skin Diseases, and the Autoimmune Bullous Skin Disease Task Force of the European Academy of Dermatology and Venereology

The Pemphigus Disease Area Index (PDAI) and Autoimmune Bullous Skin Disorder Intensity-Score (ABSIS) scores have been proposed to provide an objective measure of pemphigus activity. These scores have been evaluated only on already treated patients mainly with mild to moderate activity. The objective was to assess the interrater reliability of ABSIS and PDAI scores and their correlation with other severity markers in a large international study.

Consecutive patients with newly diagnosed pemphigus were enrolled in 31 centers. Severity scores were recorded during a 24-month period by the same two blinded investigators. Serum was collected at each visit for ELISA measurement of anti-desmoglein antibodies. The intraclass correlation coefficient (ICC) and Spearman rank correlation coefficient were calculated. A total of 116 patients with pemphigus vulgaris (n = 84) or pemphigus foliaceus (n = 32) were

[1]Department of Dermatology, Rouen University Hospital, and INSERM U 1234, Centre de référence des maladies bulleuses autoimmunes, Normandie University, Rouen, France; [2]Department of Biostatistics and Clinical Research, INSERM U 1219, Rouen University Hospital, University of Rouen, Rouen, France; [3]Department of Dermatology, Lille University Hospital, Lille, France; [4]Department of Dermatology, Universitätsklinik für Dermatologie Inselspital, Bern, Switzerland; [5]Dermatology Unit, Università di Parma, Parma, Italy; [6]Policlinico Gemelli–Università Cattolica del "Sacro Cuore," Rome, Italy; [7]Laboratory of Molecular and Cell Biology, Istituto Dermopatico dell'Immacolata, Istituto di Ricovero e Cura a Carattere Scientifico, Rome, Italy; [8]Department of Dermatology and Venereology, University Hospital Center Zagreb, University of Zagreb School of Medicine, Zagreb, Croatia; [9]Comprehensive Center for Inflammation Medicine, Department of Dermatology, University of Luebeck Ratzeburger, Luebeck, Germany; [10]Universitäts-Hautklinik Tübingen, Tübingen, Germany; [11]Department of Dermatology, Medical Center, University of Freiburg, Freiburg, Germany; [12]Department of Dermatology, Bichat University Hospital, (Assistance Publique - Hôpitaux de Paris), Paris, France; [13]Department of Dermatology, Avicenne Hospital, University Paris 13, Bobigny, France; [14]Department of Dermatology, Henri Mondor University Hospital, (Assistance Publique - Hôpitaux de Paris), Créteil, France; [15]Department of Dermatology, Saint Louis Hospital, (Assistance Publique - Hôpitaux de Paris), Paris, France; [16]Department of Dermatology, Cochin Hospital, (Assistance Publique - Hôpitaux de Paris), Paris, France; [17]Department of Dermatology, Dijon University Hospital, Dijon, France; [18]Department of Dermatology, Saint Etienne University Hospital, Saint Etienne, France; [19]Aix-Marseille University, UMR 911, INSERM CRO2, "Centre de recherche en oncologie biologique et oncophamacologie," Department of Dermatology, Hôpital Timone, Assistance Publique des Hôpitaux de Marseille, Marseille, France; [20]Department of Dermatology, Lille University Hospital, Lille, France; [21]Department of Dermatology, Rennes University Hospital, Rennes, France; [22]Department of Dermatology, Clermond-Ferrand University Hospital, Clermont-Ferrand, France; [23]Department of Dermatology, Nantes University Hospital, Nantes, France; [24]Department of Dermatology, Valence Hospital, Valence, France; [25]Department of Dermatology, Toulouse University and INSERM U1056, Toulouse, France; [26]Department of Dermatology, Bordeaux University Hospital, Bordeaux, France; [27]Department of Dermatology, Angers University Hospital, Angers, France; [28]Department of Dermatology, Limoges University Hospital, Limoges, France; [29]Department of Dermatology, Reims University Hospital, Reims, France; [30]Department of Dermatology, Tours University Hospital, Tours, France; [31]Department of Dermatology, Le Mans Hospital, Le Mans, France; [32]Department of Dermatology, Lyon University Hospital, Lyon, France; [33]Department of Dermatology, Brest University Hospital, Brest, France; [34]Department of Dermatology, Montpellier University Hospital, Montpellier, France; [35]Department of Dermatology, University of Pennsylvania, and Philadelphia Veterans Affairs Medical Center, Philadelphia, Pennsylvania, USA; [36]Department of Dermatology, St George Hospital, University of NSW, Sydney, Australia; and [37]Department of Dermatology, Philipps University Marburg, Marburg, Germany

Correspondence: Vivien Hébert, Department of Dermatology, Rouen University Hospital, 1 rue de Germont, 76031 Rouen, France. E-mail: Vivien.hebert@chu-rouen.fr

Abbreviations: ABSIS, Autoimmune Bullous Skin Disorder Intensity Score; CI, confidence interval; DLQI, Dermatology Quality of Life Index; ICC, intraclass correlation coefficient; M, month; PDAI, Pemphigus Disease Area Index; PF, pemphigus foliaceus; PGA, Physician Global Assessment; PV, pemphigus vulgaris

Received 6 December 2017; revised 20 April 2018; accepted 25 April 2018; corrected proof published online 6 October 2018

included. At baseline, the ABSIS and PDAI ICCs were 0.90 (95% confidence interval [CI] = 0.85−0.93), and 0.91(95% CI = 0.87−0.94), respectively. The ICCs for PDAI were higher in moderate and extensive pemphigus (ICC = 0.82, 95% CI = 0.63−0.92 and ICC = 0.80, 95% CI = 0.62−0.90, respectively) than in patients with intermediate (significant) extent (ICC = 0.50, 95% CI = 0.27−0.68). Conversely, the ICCs for ABSIS were lower in patients with moderate extent (ICC = 0.44, 95% CI = 0.004−0.74) than in those with intermediate or extensive forms, (ICC = 0.69, 95% CI = 0.51−0.81 and ICC = 0.75, 95% CI = 0.51−0.88, respectively). During patients' follow-up, the ICCs of both ABSIS and PDAI scores remained higher than 0.70.
ABSIS and PDAI skin (r = 0.71 and r = 0.75) but not mucosal (r = 0.32 and r = 0.37) subscores were correlated with the evolution of anti-DSG1 and anti-DSG3 ELISA values, respectively. ABSIS and PDAI scores are robust tools to accurately assess pemphigus activity.

Journal of Investigative Dermatology (2019) **139,** 31−37; doi:10.1016/j.jid.2018.04.042

INTRODUCTION

The precise assessment of skin and especially of mucosal lesions is difficult in pemphigus, a rare and severe autoimmune blistering disease. Pemphigus is caused by the production of auto-antibodies directed against desmogleins, which are desmosomal proteins responsible for the adhesion of keratinocytes (Amagai et al. 1991; Anhalt et al. 1982; Chee and Murrell, 2011; Langan et al., 2008; Sebaratnam et al., 2012; Stanley and Amagai, 2006). Pemphigus may severely impair quality of life (Chee and Murrell, 2011; Sebaratnam et al., 2012; Tabolli et al., 2014). Many severity scores have been proposed to assess pemphigus severity but none of them is widely used, because these scores did not result from an international consensus of experts (Martin and Murrell, 2006).

However, the precise assessment of the extent of skin and mucosal lesions in pemphigus patients is a major goal both for clinicians to adapt treatment and for investigators to robustly evaluate new treatment options in clinical trials. Two severity scores, the Autoimmune Bullous Skin Disorder Intensity Score (ABSIS) and the Pemphigus Disease Area Index (PDAI), have been proposed to evaluate the extent of pemphigus lesions (Pfütze et al., 2007; Rosenbach et al., 2009). The interrater reliability of ABSIS and PDAI scores has been evaluated in two preliminary studies that mainly included patients with mild to moderate pemphigus, many of whom were already treated and were assessed only once, with no follow-up evaluations (Murrell et al., 2008; Rahbar et al., 2014). As stated in an editorial in the *Journal of Investigative Dermatology,* which presented one of these studies, "The assessment of large, well-defined populations with a formal evaluation, including the usual successive methodological steps, is needed before the PDAI and ABSIS scores should be used as a research instrument" (Bastuji-Garin and Sbidian, 2009).

Therefore, we conducted an international multicenter study to prospectively assess the interrater reliability of ABSIS and PDAI scores in a large population of 116 newly diagnosed pemphigus patients with a wide spectrum of disease activity (mild/moderate, significant, and extensive) who were followed for up to 2 years. In particular, we aimed to fulfill all the methodological criteria that were detailed in the editorial to thoroughly investigate the validity of ABSIS and PDAI scoring systems in pemphigus. We also assessed the correlation between these severity scores and other clinical and biological markers of severity, including Physician Global Assessment (PGA) score, Dermatology Quality of Life Index (DLQI), and ELISA values of serum anti-DSG1 and anti-DSG3 antibodies at baseline and after the start of treatment, to measure the improvement of lesions; we also measured these in relapsing patients to assess the worsening of skin and/or mucosal lesions.

OBJECTIVES

The primary objective was to assess the interrater reliability of ABSIS and PDAI scores at baseline, first on the whole population of patients and then on subpopulations depending on (i) type of involvement (skin, using the ABSIS skin and PDAI skin subscores, or mucosal, using the ABSIS mucosa and PDAI mucosa subscores) and (ii) pemphigus extent (moderate, significant, or extensive).

Secondary objectives were (i) to assess the interrater reliability of the ABSIS and PDAI scores during patients' follow-up, (ii) to assess the correlation of these scores with the PGA and DLQI scores and ELISA values of serum anti-DSG1 and anti-DSG3 antibodies, and (iii) to assess the time to complete the scores.

RESULTS

Baseline characteristics of patients

Overall, 116 patients (68 women and 48 men) were enrolled in the study. Thirty-two patients had pemphigus foliaceus (PF), and 84 had pemphigus vulgaris (PV) (23 with exclusive mucosal involvement, 32 with exclusive skin involvement, and 61 with mucosal and skin lesions). Mean age was 53 years (standard deviation = 14.9 years, range = 19−84 years).

At baseline, median (range) ABSIS and PDAI activity scores of the whole population were 37 out of 206 points (0.5−124.5 points) and 27.5 out of 250 points (3−114 points), respectively. Median (range) PGA score was 6 out of 10 points (1−10 points), and median DLQI score was 8 out of 30 points (0−29 points). The number of patients with PV or PF; moderate, significant or extensive pemphigus; and median corresponding PDAI, ABSIS, PGA, and DLQI scores and anti-DSG1 and anti-DSG3 antibody ELISA values are shown in Table 1.

Interrater reliability

At baseline, a high interrater reliability was observed for both the ABSIS (intraclass correlation coefficient [ICC] = 0.90; 95% confidence interval [CI] = 0.85−0.93) and PDAI scores (ICC = 0.91, 95% CI = 0.87−0.94) and, to a lesser degree, for the PGA score (ICC = 0.80, 95% CI = 0.72−0.86).

According to pemphigus extent, the baseline ICC values for PDAI were significantly higher in patients with mild/moderate (ICC = 0.82; 95% CI = 0.63−0.92) and extensive (ICC = 0.80, 95% CI = 0.62−0.90) extent than in patients with

Table 1. Baseline clinical characteristics and desmoglein ELISA values of pemphigus patients

Characteristics	ABSIS[1], Median (Range)	PDAI[1], Median (Range)	PGA[1], Median (Range)	Anti-DSG1[1] in IU/ml, Median (Range)	Anti-DSG3[1] in IU/ml, Median (Range)
Whole population (n = 116)	36.8 (0.5−124.5)	27 (3−114)	6 (1−10)	129 (1− 2,065)	240 (1− 5,217)
Pemphigus type					
Pemphigus vulgaris (n = 84)	39 (0.5−106)	29 (4−114)	6 (1−10)	52 (1−2,065)	750 (1−5,217)
Only mucosal involvement (n = 23)	39.3 (6−56)	23 (5−76)	6 (3−10)	5 (1−75)	162.5 (1− 3,800)
Skin and mucosal involvement (n = 61)	36 (0.5−106)	32 (4−114)	6 (1−10)	118 (1− 2,068)	910 (1− 5,217)
Pemphigus foliaceus (n = 32)	25.3 (1−124.5)	24.5 (3−95)	6 (1−9)	640 (4− 2,800)	1 (1−179)
Pemphigus extent[1]					
Mild/moderate (n = 26)	6 (0.5−15)	10 (3−14)	4 (1−3)	6 (1−1,960)	185 (1− 4,450)
Significant (n = 60)	36 (17−52.8)	27 (15−44)	6 (4−7)	89.5 (1− 2,800)	183 (1− 5,217)
Extensive (n = 30)	57.5 (53−124.5)	64.5 (45−114)	8 (8−10)	207.5 (1− 2,320)	622.5 (1−3,975)

Abbreviations: ABSIS, Autoimmune Bullous Skin Disorder Intensity Score; IU, international units; PDAI, Pemphigus Disease Area Index; PGA, Physician Global Assessment.
[1]According to the thresholds proposed by Boulard et al. (2016).

significant extent (ICC = 0.50, 95% CI = 0.27−0.68), with $P = 0.017$ and $P = 0.022$, respectively. ABSIS ICC values were borderline significantly higher in patients with significant (ICC = 0.69, 95% CI = 0.51−0.81) and extensive pemphigus (ICC = 0.75, 95% CI = 0.51−0.88) than in those with mild/moderate pemphigus (ICC = 0.44, 95% CI = 0.004−0.74), with $P = 0.178$ and $P = 0.116$, respectively. Finally, ICC (95% CI) values for PGA in mild/moderate, significant, and extensive pemphigus were 0.51 (0.07−0.83), 0.56 (0.37−0.71), and 0.65 (0.35−0.86), respectively.

According to type of involvement (skin or mucosa), the ICC values (95% CI) for PDAI skin and ABSIS skin were 0.97 (0.96−0.98) and 0.96 (0.94−0.97), respectively, and those for PDAI mucosa and ABSIS mucosa were 0.91 (0.87−0.94) and 0.96 (0.94−0.97), respectively. In cases of discrepancy between investigators, the mean (± standard deviation) interrater difference was higher for assessment of mucosal lesions than for skin lesions: PDAI mucosa, 7.5 ± 10 points versus PDAI skin, 4.5 ± 4.3 points, and ABSIS mucosa, 6.7 ± 6.7 points versus ABSIS skin, 4.9 ± 6.6 points.

Figure 1 shows the evolution of the PDAI, ABSIS, and PGA scores during the 2-year follow-up. Figure 2 shows that the ABSIS and PDAI ICCs remained higher than 0.70 during follow-up. In particular, the ICC values (95% CI) for PDAI and ABSIS at month (M) 1 were 0.84 (0.77−0.89) and 0.90 (0.88−0.95), respectively, and those at M3 were 0.70 (0.58−0.79) and 0.91 (0.86−0.94), respectively. The slightly lower ICC values observed at M3 for PDAI score corresponded to an ICC (95% CI) of 0.72 (0.61−0.80) for PDAI skin and of 0.87 (0.81−0.91) for PDAI mucosa.

Correlation between severity scores and other markers of disease severity

Baseline correlations. At baseline, Spearman coefficient correlation was $r = 0.57$ ($P < 0.0001$) between ABSIS and PDAI scores, $r = 0.68$ ($P < 0.0001$) between PDAI and PGA scores, and $r = 0.60$ ($P < 0.0001$) between ABSIS and PGA scores. ABSIS skin and PDAI skin subscores were highly correlated ($r = 0.87$, $P < 0.0001$), as were mucosal subscores ($r = 0.85$, $P < 0.0001$) (Figure 3).

At baseline, PDAI skin and ABSIS skin subscores were highly correlated with anti-DSG1 ELISA values: $r = 0.84$ ($P < 0.0000$) and $r = 0.77$ ($P < 0.0001$), respectively. PDAI mucosa and ABSIS mucosa subscores were also correlated with anti-DSG3 ELISA values: $r = 0.62$ ($P < 0.0001$) and 0.57 ($P < 0.0001$), respectively (Figure 3).

Baseline ABSIS and PDAI scores were both weakly correlated with the DLQI score in the whole sample: $r = 0.24$ ($P = 0.02$) and $r = 0.33$ ($P = 0.001$), respectively. According to the type of pemphigus, the PDAI score was weakly correlated with the DLQI score in PV ($r = 0.30$, $P = 0.01$) and PF patients ($r = 0.39$, $P = 0.05$). No correlation was found between ABSIS and DLQI scores in the PV and PF subpopulations.

Correlations during disease course. Overall, 100 of 116 patients achieved complete remission (ABSIS and PDAI scores = 0) at any time during the study, including 18 patients (15.5%) at the M1 evaluation and 39 (33.6%) and 68 (58.6%) patients at the M3 and M6 evaluations, respectively.

Among these latter patients, anti-DSG1 and anti-DSG3 antibodies were observed in 8 (11.7%) and 23 patients (33.8%), with mean ELISA values of 113.6 ± 62.9 IU/ml (anti-

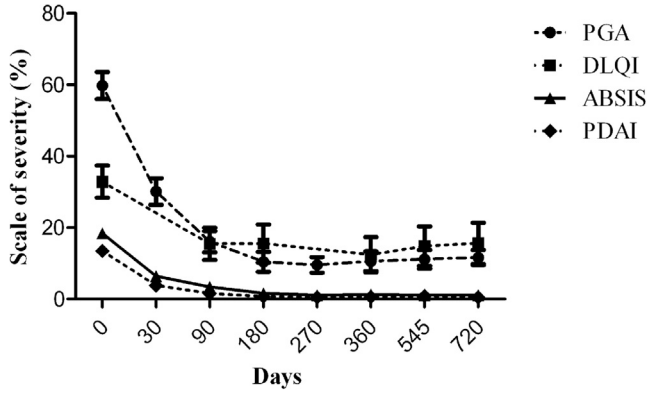

Figure 1. Evolution of ABSIS, PDAI, and PGA severity scores and DLQI quality of life score during patients' follow-up. The vertical bars correspond to the 95% confidence intervals. ABSIS, Autoimmune Bullous Skin Disorder Intensity Score; DLQI, Dermatology Quality of Life Index; PDAI, Pemphigus Disease Area Index; PGA, Physician Global Assessment.

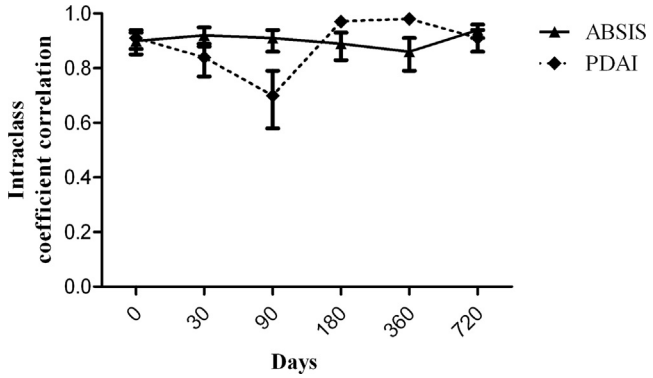

Figure 2. Interrater reliability of ABSIS and PDAI scores (intraclass coefficient correlation) during patients' follow-up. The vertical bars correspond to the 95% confidence intervals. ABSIS, Autoimmune Bullous Skin Disorder Intensity Score; PDAI, Pemphigus Disease Area Index.

DSG1 antibodies) and 152.5 ± 235.8 IU/ml (anti-DSG3 antibodies), respectively. More precisely, 6 patients had anti-DSG1antibodies alone, 21 had anti-DSG3 antibodies alone, and 2 had both anti-DSG1 and anti-DSG3 antibodies.

During the follow-up, ABSIS and PDAI scores remained correlated ($r = 0.55$, $P < 0.0001$). Both ABSIS and PDAI scores remained correlated with PGA: $r = 0.45$ ($P < 0.0001$) and $r = 0.54$ ($P < 0.0001$), respectively; they were no longer correlated with the DLQI score. We then evaluated the correlation between the ABSIS and PDAI scores during the first 3-month period after the start of treatment to know whether these scores are useful for evaluating the initial improvement of patients' condition and the effect of treatment. From baseline to M1, ABSIS and PDAI scores remained correlated ($r = 0.54$, $P < 0.0001$), as they did for the M1 to the M3 evaluations ($r = 0.68$, $P < 0.0001$). Table 2 shows the correlation between ABSIS and PDAI scores during this initial period, depending on the mild, moderate, or severe type of pemphigus. Apart from a poor correlation ($r = 0.2$) from baseline to M1 in patients with mild types of pemphigus, the correlation between ABSIS and PDAI scores was between 0.44 and 0.78 for moderate and severe pemphigus subtypes during this initial 3-month period of treatment.

The absolute improvement of the PDAI skin and ABSIS skin activity subscores was highly correlated with the absolute decrease of anti-DSG1 antibodies: $r = 0.75$ ($P < 0.0001$) and $r = 0.71$ ($P < 0.0001$), respectively. Conversely, only a weak correlation was observed between the absolute change in the PDAI mucosa and ABSIS mucosa subscores and the absolute change in anti-DSG3 antibodies: $r = 0.37$ ($P < 0.001$) and $r = 0.32$ ($P = 0.003$), respectively. To further assess the correlation between severity scores and anti-DSG antibodies during disease course, we studied the correlation between the ABSIS, PDAI, and their respective skin and mucosal subscores with the evolution of anti-DSG1 and anti-DSG 3 ELISA values in the 17 patients who relapsed during the 2 years of follow-up. Because patients relapsed at different time points during follow-up, we compared the evolution of the ABSIS and PDAI scores and that of anti-DSG1 and anti-DSG3 antibodies between the last evaluation and serum collection performed during the period of complete remission (before relapse) and the next evaluation performed during skin and/or mucosal relapse. Whereas the ABSIS and PDAI scores remained correlated ($r = 0.64$, $P = 0.0061$), we did not observe a correlation between the PDAI skin and ABSIS skin and anti-DSG1 antibodies ($r = 0.28$, $P = 0.2841$ and $r = 0.41$, $P = 0.1343$, respectively), nor between the PDAI

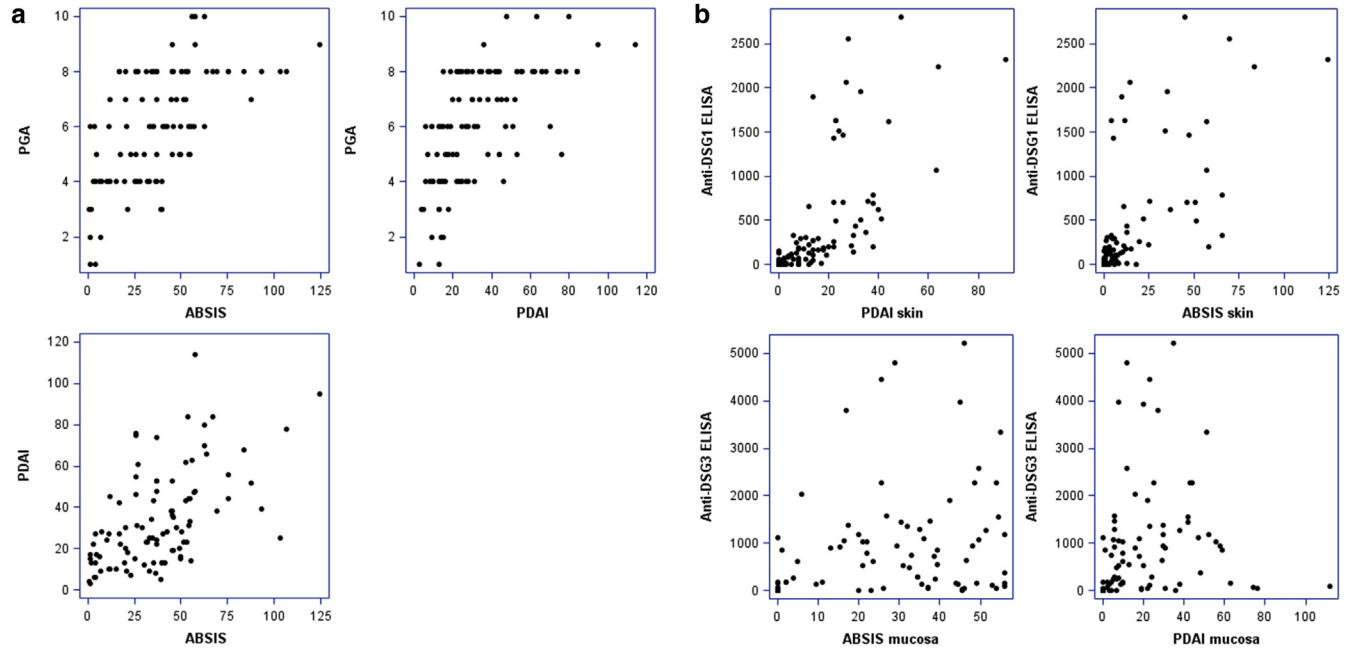

Figure 3. Baseline correlations of total scores and skin or mucosa ABSIS and PDAI subscores with PGA and anti-DSG ELISA values. Scatter diagram depicting the baseline correlation between (**a**) ABSIS, PDAI, and PGA severity scores and (**b**) PDAI skin and ABSIS skin and PDAI mucosa and ABSIS mucosa subscores with anti-DSG1 and anti-DSG3 ELISA values, respectively. ABSIS, Autoimmune Bullous Skin Disorder Intensity Score; PDAI, Pemphigus Disease Area Index; PGA, Physician Global Assessment.

Table 2. Correlations between ABSIS and PDAI scores from baseline to M1 and M3 according to the mild, moderate, or severe types of pemphigus

Type of Pemphigus	Baseline to M1	M1 to M3
Mild types		
Spearman rank correlation coefficient	0.1986	0.6305
95% CI	−0.3258 to 0.6296	0.1805 to 0.8623
P	0.4447	0.0088
Moderate types		
Spearman rank correlation coefficient	0.4371	0.7795
95% CI	0.1218 to 0.6722	0.5846 to 0.8894
P	0.0068	<0.0001
Severe types		
Spearman rank correlation coefficient	0.6134	0.4843
95% CI	0.1533 to 0.8549	−0.05398 to 0.8045
P	0.0115	0.0673

Abbreviation: CI, confidence interval; M, month.

mucosa and ABSIS mucosa and anti-DSG3 antibodies ($r = 0.24$, $P = 0.3636$ and $r = 0.27$, $P = 0.3014$, respectively).

Time for instrument completion

The mean time for carrying out the PDAI was 5.2 minutes. It was progressively quicker, from 8.3 minutes at baseline to 3.6 minutes at month 24. Nine percent of investigators found the PDAI easy to use, 21% quite easy, 35% practical, 24% difficult, and 11% imprecise or too long. The mean time for ABSIS completion was 5.4 minutes, and this also decreased over time from 8.1 minutes at baseline to 3.9 minutes at month 24. Six percent of investigators found the ABSIS easy to use, 13% quite easy, 31% practical, 39% difficult, and 11% imprecise or too long.

DISCUSSION

Our results definitely validate the ABSIS and PDAI scores, showing an excellent interrater reliability for both scores (ICC ≥ 0.90) and a good correlation between these two scores ($= 0.55$). The inter-rater reliability of the ABSIS was high in scoring patients with significant or extensive pemphigus (ICC = 0.69 and ICC = 0.75, respectively), whereas it was lower in scoring patients with mild/moderate extent (ICC = 0.44). This is likely because the ABSIS uses the rule of nine to estimate body surface involvement, which makes the scoring of patients with limited pemphigus extent difficult. The PDAI score had a high interrater reliability for scoring patients with mild/moderate and extensive pemphigus (ICC = 0.82 and ICC = 0.80, respectively) and a lower reproducibility for assessing patients with intermediate extent (ICC = 0.50). This is a quite common weakness in many scoring systems, such as the Psoriasis Area Severity Index in psoriasis, which is due to the heterogeneity of patients with intermediate disease extent (Paul et al., 2010). It may seem surprising that ICC estimates were lower in subgroups than for the overall sample: for example, 0.91 for the PDAI ICC for the overall sample compared with 0.82, 0.50, and 0.80 for mild/moderate, significant, and extensive forms, respectively. However, this was expected, because ICC estimates are known to be lower when observations are more homogeneous, which is the case in subgroups compared with the overall sample. In case of discrepancy, the difference between investigators was rather low for the assessment of skin lesions (4.5 to 5 points) and slightly higher for the assessment of mucosal lesions (6.7 to 7.5 points).

Our results show a higher reproducibility for both PDAI and ABSIS scores than was reported by Rosenbach et al. (2009) in a study of 15 patients (0.76 and 0.77, respectively) and a lower reproducibility than was reported by Rahbar et al. (2014) in a cross-sectional study (0.98 and 0.97, respectively). These two studies mainly included already-treated patients, most of whom had low to moderate disease activity as a consequence and who had been evaluated once only with no follow-up.

The ABSIS and PDAI scores were not only reliable instruments to measure pemphigus extent at baseline, but they also showed a high reproducibility during follow-up, both to measure the improvement of lesions after the start of treatment and also to assess the worsening of skin and/or mucosal lesions in relapsing patients. This feature is particularly important for assessing the evolution of skin and/or mucosal lesions under treatment, because as previously reported, only anti-DSG1, but not anti-DSG3, antibodies seem useful for following the course of pemphigus patients (Abasq et al., 2009; Patsatsi et al., 2014). Indeed, this study showed that anti-DSG3 antibodies were weakly correlated with the evolution of mucosal lesions over time, suggesting that the PDAI and ABSIS mucosa subscores are more accurate than anti-DSG3 ELISA values for assessing the evolution of mucosal lesions and adapting treatment. In particular, whereas the correlation between ABSIS and PDAI scores in relapsing patients ($r = 0.64$) remained in the same order of magnitude as at baseline ($r = 0.55$), these scores were only weakly correlated with the evolution of anti-DSG antibodies under treatment, reinforcing the usefulness of the ABSIS and PDAI scores in the follow-up of patients.

The evolution of the PGA score was also correlated with the ABSIS and PDAI scores. However, the PGA score is based on a physician's subjective impression, with no clear definition of what the different marks from 0 to 10 correspond to. Additionally, it does not allow separate assessment of the evolution of skin and mucosal lesions. Similarly, the evolution of the DLQI score was not correlated with the evolution of pemphigus lesions and cannot be used as a tool to adapt treatment in pemphigus patients. The recently published Autoimmune Bullous Disease Quality of Life and Treatment of Autoimmune Bullous Disease Quality of Life scores might be better correlated with disease activity (Tjokrowidjaja et al., 2013).

Finally, these scores are feasible in clinical practice, taking an average of approximately 5 minutes, which is a little bit longer than the time previously reported (Rahbar et al., 2014). Investigators generally considered the PDAI easier to use than the ABSIS.

This study has several major strengths. First, it was performed prospectively on a large, multicenter, international cohort of pemphigus patients who were followed up for 2 years. Because it only included newly diagnosed pemphigus cases not yet treated, this study allowed assessment of these

scores not only in patients with mild to moderate pemphigus activity but also in patients with significant and extensive pemphigus. It fulfills all the methodological criteria that were recently detailed as necessary to thoroughly investigate the validity of the ABSIS and PDAI scoring systems in pemphigus (Bastuji-Garin and Sbidian, 2009).

Selection bias is unlikely in this study, because it included consecutive patients recruited both in secondary and tertiary care centers with mild/moderate, significant, or extensive pemphigus. The M1 and M3 dates for early evaluation of these scores were chosen by a panel of international experts, because these dates correspond to common time points used to evaluate disease activity and the effect of treatment in most clinical trials on pemphigus.

Few patients (15.5% and 33.6%, respectively) were in complete remission at these dates, thus allowing a valid assessment of these scores in patients who still had active disease. Despite the extremely high interrater reliability, evidenced both at baseline and at the M1 and M3 evaluations, weekly evaluations would have reinforced the validity of our results.

Overall, this study provides strong evidence that the ABSIS and PDAI scores are robust tools to accurately assess pemphigus activity, both at the time of diagnosis and during disease course, which is of major interest for clinicians to adapt treatment. The high interrater reliability of these scoring systems will allow valuable intergroup comparisons of disease activity in randomized clinical trials. Finally, one might question which scoring system investigators should choose. We found that the PDAI score had a higher reproducibility than the ABSIS in scoring patients with mild pemphigus. Conversely, it had a lower reproducibility than the ABSIS in assessing patients with intermediate extent. Overall, 30% and 19% of investigators found the PDAI and ABSIS scores, respectively, easy or quite easy to use. Finally, the PDAI has been developed by a consensus of international experts, whereas the ABSIS was proposed by German investigators. Further research is needed to assess the prognostic value of these scores.

PATIENTS AND METHODS

Study Population

We conducted a prospective, international, multicenter study in 31 French, German, Italian, Swiss, and Croatian departments of dermatology (secondary and tertiary care centers) between July 2009 and May 2015. Consecutive patients aged 18 years or older with newly diagnosed pemphigus were included. Diagnosis of either PV or PF was based on (i) characteristic clinical features; (ii) histological analysis of a skin or mucosal biopsy showing acantholysis, intraepithelial blistering, or eosinophilic spongiosis; (iii) direct immunofluorescence examination showing IgG and/or C3 deposits on keratinocyte cell membrane; and (iv) detection of circulating anti-DSG1 and/or anti-DSG3 autoantibodies by ELISA assays (Amagai et al., 1999; Kasperkiewicz et al., 2017). Treatments were not controlled in this study. They varied among countries and mainly consisted of oral corticosteroids alone or associated with conventional immunosuppressants or rituximab in some patients.

All patients gave signed informed consent before inclusion. The study was approved by the corresponding local ethics committee.

Assessment of disease extent

Disease extent was evaluated at baseline and during the follow-up visits at M1, M3, M6, M12, and M24 by the same two investigators blinded to the results of each other and using ABSIS, PDAI, and PGA scores. All investigators were dermatologists with extensive experience in the diagnosis and treatment of pemphigus patients.

The ABSIS score of skin involvement is based on the extent of the body surface area assessed using Wallace's "rule of nines" and type of skin lesions (Livingston et al., 2000). The body surface area value is then multiplied by an index reflecting the predominant lesions: 1.5 (erosive exudative lesions, bullae, or Nikolsky sign positivity), 1.0 (erosive dry lesions), or 0.5 (re-epithelialized lesions). ABSIS oral involvement is evaluated by scoring 11 mucosal sites by 1 (presence of lesions) or 0 (absence of lesions) and by completing a subjective severity scale based on discomfort during eating and drinking. The ABSIS ranges from 0 to 206 points, with 150 points for skin involvement, 11 points for oral involvement, and 45 points for subjective discomfort; higher scores denote worse disease. A score between 0 and 16 corresponds to moderate pemphigus, between 17 and 52 to significant pemphigus, and higher than 52 to extensive pemphigus (Boulard et al., 2016).

The PDAI has a score ranging from 0 to 263 points, with 250 points representing disease activity (120, 10, and 120 points for skin, scalp, and mucosal activity, respectively) and 13 points representing disease damage. However, the damage component was not included in our analysis. For skin activity assessment, 12 anatomic sites are assigned a score according to disease extent: 0 (no lesions), 1 (1−3 lesions, up to 1 lesion $>$ 2 cm in any diameter, all $\leq$ 6 cm), 2 (2−3 lesions, at least 2 lesions $>$ 2 cm, all $\leq$ 6 cm), 3 ($>$ 3 lesions, all $\leq$ 6 cm), 5 ($>$ 3 lesions and/or 1 lesion $>$ 6 cm), or 10 ($>$ 3 lesions and/or at least 1 lesion $>$ 16 cm or entire area affected). Scalp activity is assigned a score based on the presence of blisters, erosions, or erythema of 0 (no activity), 1 (one quadrant affected), 2 (two quadrants affected), 3 (3 quadrants affected), 4 (whole skull affected), or 10 (at least 1 lesion $>$ 6 cm). For mucosal activity assessment, 12 mucosal sites are assigned a score based on the presence of erosions or blisters: 0 (absent), 1 (1 lesion), 2 (2−3 lesions), 5 ($>$ 3 lesions or 2 lesions $>$ 2 cm), or 10 (entire area). A score between 0 and 14 corresponds to moderate pemphigus, between 15 and 44 to significant pemphigus, and higher than 44 to extensive pemphigus (Boulard et al., 2016).

The PGA is a visual analog 10-point scale, based on a physician's subjective impression from 0 (no lesions) to 10 (worst skin and mucosae condition imaginable). It has been used in clinical trials because it is fast and easy to use (Tabolli et al., 2014).

Patients' quality of life was evaluated by the DLQI translated into different languages (Kasperkiewicz et al., 2017). It includes 10 questions, with a total score between 0 and 30. The DLQI was used because the Autoimmune Bullous Disease Quality of Life and Treatment of Autoimmune Bullous Disease Quality of Life were not available at the time the study was designed.

Serum samples were collected at each visit, stored on site, and centrally analyzed at the end of the study in the Immunology Laboratory of Rouen University Hospital for measurement of anti-DSG1 and anti-DSG3 antibody ELISA values, using commercially available ELISA-DSG1 and ELISA-DSG3 assays (Euroimmun, Lübeck, Germany). ELISA values higher or equal to 20 IU/ml were considered positive. Beyond 200 IU/ml, which corresponds to the upper limit of the assay, additional dilutions were performed.

Statistical analysis

Scores were prospectively recorded on standardized forms. Patients with more than one missing score out of four were excluded from the analysis. The target sample size (n = 100) was calculated for the primary objective of this study, which was to assess the interrater reliability through estimation of the ICC with good precision, as measured by the width of the 95% ICC CI. With n = 100, the expected width was ±0.15 for an ICC of 0.5 and ±0.07 for an ICC of 0.8. These figures translated into respective widths of ±0.21 and ±0.10 for half the overall sample size (n = 50), that is, the expected size of the subgroup with significant extent, and into respective widths of ±0.31 and ±0.14 for one quarter of the overall sample size (n = 25), that is, the expected size of the subgroups with mild/moderate or extensive extent.

Interrater reliability was assessed by estimating the ICC overall and according to severity (mild/moderate, significant, extensive). This was done at each study visit for the overall ABSIS and PDAI scores, their respective skin and mucosal subscores (ABSIS skin and PDAI skin; ABSIS mucosa and PDAI mucosa), and per severity subgroups (mild/moderate, significant, extensive).

Correlations between ABSIS and PDAI scores and with other severity markers (PGA, DLQI, and anti-DSG1 and anti-DSG3 antibody ELISA) were assessed using Spearman rank correlation coefficient. Correlations were assessed for baseline values of these scores or other severity markers and for their absolute changes between time 0 and M6. Correlations were not calculated after 6 months of follow-up, because severity scores were returned to 0 or almost 0 in most patients. SAS, version 9.3 (SAS institute, Cary, NC) and GraphPad Prism, version 5.0 (GraphPad Software, San Diego, CA) software were used.

ORCID

Pierre Vabres: http://orcid.org/0000-0001-8693-3183

ACKNOWLEDGMENTS

The study was supported by grants from the Seventh Framework Programme of the European Union (FP7-HEALTH-2007-2.4.4-1), natural course and pathophysiology of rare diseases, and the French Society of Dermatology.

REFERENCES

Abasq C, Mouquet H, Gilbert D, Tron F, Grassi V, Musette P, et al. ELISA testing of anti-desmoglein 1 and 3 antibodies in the management of pemphigus. Arch Dermatol 2009;145:529–35.

Amagai M, Klaus-Kovtun V, Stanley JR. Autoantibodies against a novel epithelial cadherin in pemphigus vulgaris, a disease of cell adhesion. Cell 1991;67:869–77.

Amagai M, Komai A, Hashimoto T, Shirakata Y, Hashimoto K, Yamada T, et al. Usefulness of enzyme-linked immunosorbent assay using recombinant desmogleins 1 and 3 for serodiagnosis of pemphigus. Br J Dermatol 1999;140:351–7.

Anhalt GJ, Labib RS, Voorhees JJ, Beals TF, Diaz LA. Induction of pemphigus in neonatal mice by passive transfer of IgG from patients with the disease. N Engl J Med 1982;306:1189–96.

Bastuji-Garin S, Sbidian E. How to validate outcome instruments for pemphigus. J Invest Dermatol 2009;129:2328–30.

Boulard C, Duvert Lehembre S, Picard-Dahan C, Kern JS, Zambruno G, Feliciani C, et al. Calculation of cut-off values based on the ABSIS and PDAI pemphigus scoring systems for defining moderate, significant and extensive types of pemphigus. Br J Dermatol 2016;175:142–9.

Chee S-N, Murrell DF. Pemphigus and quality of life. Dermatol Clin 2011;29:521–5.

Kasperkiewicz M, Ellebrecht CT, Takahashi H, Yamagami J, Zillikens D, Payne AS, et al. Pemphigus. Nat Rev Dis Primers 2017;3:nrdp201726.

Langan SM, Smeeth L, Hubbard R, Fleming KM, Smith CJP, West J. Bullous pemphigoid and pemphigus vulgaris—incidence and mortality in the UK: population based cohort study. BMJ 2008;337:a180.

Livingston EH, Lee S. Percentage of burned body surface area determination in obese and nonobese patients. J Surg Res 2000;91:106–10.

Martin L, Murrell DF. Measuring the immeasurable: a systematic review of outcome measures in pemphigus. Australas J Dermatol 2006;47:A32–3.

Murrell DF, Dick S, Ahmed AR, Amagai M, Barnadas MA, Borradori L, et al. Consensus statement on definitions of disease, end points, and therapeutic response for pemphigus. J Am Acad Dermatol 2008;58:1043–6.

Patsatsi A, Kyriakou A, Giannakou A, Pavlitou-Tsiontsi A, Lambropoulos A, Sotiriadis D. Clinical significance of anti-desmoglein-1 and -3 circulating autoantibodies in pemphigus patients measured by Area Index and Intensity score. Acta Derm Venereol 2014;94:203–6.

Paul C, Gourraud P-A, Bronsard V, Prey S, Puzenat E, Aractingi S, et al. Evidence-based recommendations to assess psoriasis severity: systematic literature review and expert opinion of a panel of dermatologists. J Eur Acad Dermatol Venereol 2010;24(Suppl. 2):2–9.

Pfütze M, Niedermeier A, Hertl M, Eming R. Introducing a novel Autoimmune Bullous Skin Disorder Intensity Score (ABSIS) in pemphigus. Eur J Dermatol 2007;17:4–11.

Rahbar Z, Daneshpazhooh M, Mirshams-Shahshahani M, Esmaili N, Heidari K, Aghazadeh N, et al. Pemphigus disease activity measurements: Pemphigus Disease Area Index, Autoimmune Bullous Skin Disorder Intensity Score, and Pemphigus Vulgaris Activity Score. JAMA Dermatol 2014;150:266–72.

Rosenbach M, Murrell DF, Bystryn J-C, Dulay S, Dick S, Fakharzadeh S, et al. Reliability and convergent validity of two outcome instruments for pemphigus. J Invest Dermatol 2009;129:2404–10.

Sebaratnam DF, Frew JW, Davatchi F, Murrell DF. Quality-of-life measurement in blistering diseases. Dermatol Clin 2012;30:301–7.

Stanley JR, Amagai M. Pemphigus, bullous impetigo, and the staphylococcal scalded-skin syndrome. N Engl J Med 2006;355:1800–10.

Tabolli S, Pagliarello C, Paradisi A, Cianchini G, Giannantoni P, Abeni D. Burden of disease during quiescent periods in patients with pemphigus. Br J Dermatol 2014;170:1087–91.

Tjokrowidjaja A, Daniel BS, Frew JW, Sebaratnam DF, Hanna AM, Chee S, et al. The development and validation of the treatment of Autoimmune Bullous Disease Quality Of Life Questionnaire, a tool to measure the quality of life impacts of treatments used in patients with autoimmune blistering disease. Br J Dermatol 2013;169:1000–6.