Pemphigus Vulgaris Activity Score and Assessment of Convergent Validity

Cheyda Chams-Davatchi, Ziba Rahbar, Maryam Daneshpazhooh, Seyed Mohammad Ali Mortazavizadeh, Maryam Akhyani, Nafiseh Esmaili, and Kamran Balighi

Autoimmune Bullous Diseases Research Center, Tehran University of Medical Sciences, Tehran, Iran

Received: 29 May 2012; Received in revised form: 11 Dec. 2012; Accepted: 5 Jan. 2013

Abstract- Pemphigus is a rare autoimmune blistering disease with different phenotypes. The evaluation of therapeutic interventions requires a reliable, valid and feasible to use measurement. However, there is no gold standard to measure the disease activity in clinical trials. In this study we aimed to introduce the pemphigus vulgaris activity score (PVAS) measurement and to assess the convergent validity with the experts’ opinion of disease activity. In PVAS scoring, the distribution of pemphigus vulgaris antigen expression in different anatomical regions is taking into account with special consideration of the healing process. PVAS is a 0-18 scale, based on the extent of mucocutaneous involvement, type of lesion and the presence of Nikolsky’s sign. The sum of the scores of total number of lesions, number of different anatomic regions involvement and Nikolsky’s sign is weighted by the type of lesion. In the present study, PVAS was assessed in 50 patients diagnosed with pemphigus vulgaris by one dermatologist. Independently, five blinded experts scored all the patients through physician’s global assessment (PGA). The convergent validity with experts’ opinion was assessed. The Spearman coefficient of correlation showed the acceptable value of 0.751 (95%CI: 0.534-0.876). PVAS is a valid, objective and simple-to-use scoring measurement. It showed a good correlation with PGA of pemphigus disease activity in Iranian patients with pemphigus vulgaris.

© 2013 Tehran University of Medical Sciences. All rights reserved.

Keywords: Expert; Pemphigus; Physician’s Global Assessment; Validity

Introduction

Pemphigus is a chronic autoimmune bullous disease having morbidity and mortality, with multiple clinical varieties. In recent decades, much progress was achieved on the pathophysiology of the disease, and new efficient drugs were introduced that have improved the prognosis of this severe disease.

Autoimmune pemphigus consists of three phenotypically different diseases of pemphigus vulgaris (PV), pemphigus foliaceous and paraneoplastic pemphigus. Several studies on pemphigus vulgar showed that the age-of-onset, disease phenotype, severity, and also the incidence of the disease vary from one continent to another (1-3).

Multicenter randomized controlled trials studies are essentials in rare autoimmune blistering disease treatment. But to assess the therapeutic efficacy of different treatment modalities, a valid, and feasible to use clinical scoring system is required (4). In fact there is no approved gold standard for measuring disease activity in patients with PV (5) which might be related to the rarity of the disease. Evaluation of patients’ quality of life (QOL) might be used instead to assess the effectiveness of therapeutic interventions and monitor disease over time. However, it may not correlate with changes in clinical disease activity. Some studies (6,7) noted discrepancies such as relatively poorer QOL in the early limited stages of PV, or better scores in more advanced disease patients with adequate coping empowers. Although serum anti-desmoglein 1 and 3 Enzyme-linked immunosorbent assay (ELISA) titer in pemphigus could be a good predictor of disease activity, they are not routinely available. Also, recently it was shown that anti-Dsg3 autoantibody did not appropriately correlate with the PV course of the disease (8). Two clinical scoring outcome measurements of autoimmune bullous skin disorder intensity score (ABSIS) and pemphigus disease area index (PDAI) were developed, but the former accounts only oral mucosa and the latter is complicated to use (9).

In this study we presented the disease activity...
measurement that is used in the Autoimmune Bullous Diseases Research Center (ABDRC), Iran. Also we assessed the convergent validity with experts’ opinion to estimate how much it can be close to autoimmune blistering disease experts’ opinion of disease activity. Pemphigus vulgaris activity score (PVAS) is an objective, rapid and simple clinical method that would be helpful in estimating the disease activity, even for non-expert physicians.

**Materials and Methods**

First step: The PVAS was designed based on disease activity variables from the PV literature and authors practical experience. Second step: The PVAS was tested by one expert in PV, through physician global assessment (PGA) in 120 PV patients of wide disease activity range. Third step: Upon concordance of the PGA and the PVAS scoring system, validation study for contiguity with experts’ opinion of pemphigus disease activity was performed within 50 other patients that are presented here.

**Participants**

Fifty consecutive PV patients were selected to enroll in the study. The pemphigus disease diagnosis was confirmed by histopathology and direct immunofluorescence studies. All the patients signed the informed consent form.

A dermatologist calculated the PVAS score for each patient. Also, five experts in autoimmune bullous disease were asked to evaluate the patients individually and give their PGA, a visual analog scale to rate disease activity by general overall impression, for each patient from 0 to 18, minimum to maximum disease activity, respectively. They were kept blinded to PVAS results.

**The PVAS scoring system**

PVAS was based on: 1- The extent of lesions: total number of lesions and different anatomical regions in skin and all mucous membrane. 2- The presence of Nikolsky’s sign. 3- The lesion type according to the healing process (Table 1).

**Table 1.** Components of PVAS: Skin involvement [s] is presented by the sum of the number of skin lesions (N), distribution in different body areas (D) and presence of Nikolsky’s sign (S), weighed by type of the lesion (T). The skin lesion type coefficient for bulla or erosions is 1, for crusts is 0.5 and for pigmentation is 0. Mucous membrane involvement [m] is presented by the sum of the number of mucous membrane lesions (N), distribution (D), is weighed by type of the lesion (T). The mucous lesions type coefficient for bulla or erosions is 1 and for ulceration is 0.5.

<table>
<thead>
<tr>
<th>PVAS =</th>
<th>$\text{T_s} \left( \text{N}_s \times \text{D}_s \times \text{S}_s \right) + \text{T_m} \left( \text{N}_m \times \text{D}_m \right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{T}_s$:</td>
<td>Type of skin lesion</td>
</tr>
<tr>
<td>$\text{N}_s$:</td>
<td>Number of skin lesion</td>
</tr>
<tr>
<td>$\text{D}_s$:</td>
<td>Distribution of skin lesion, One point for each anatomical area</td>
</tr>
<tr>
<td>$\text{S}_s$:</td>
<td>Nikolsky’s sign: pressure induced blister</td>
</tr>
<tr>
<td>$\text{T}_m$:</td>
<td>Type of mucosal lesion</td>
</tr>
<tr>
<td>$\text{N}_m$:</td>
<td>Number of mucosal lesion</td>
</tr>
<tr>
<td>$\text{D}_m$:</td>
<td>Distribution of mucosal lesion. One point for each anatomical area</td>
</tr>
</tbody>
</table>

Maximum skin score: 11, maximum mucosa score: 7, maximum total score $[\text{T}_s(\text{N}_s \times \text{D}_s \times \text{S}_s) + \text{T}_m(\text{N}_m \times \text{D}_m)]$: 18
Pemphigus vulgaris activity score

Through the process of healing in each pemphigus flare, three clinical stages would be followed in the skin; Active blisters with mostly “exudative erosions” and positive Nikolsky’s sign progress to mostly erosive but “dry crusted” lesion that gradually improve into re-epithelialized post-inflammatory hyperpigmentation (10). On the other hand, in the mucosa, an acute onset oral PV shows rough superficial erosions of the oral mucosa that tears easily without a visible tendency to healing (10). When the erosion starts to heal, it changes the aspect and improves into ulcer that the borders become more well-defined, paradoxically the lesion seems deep and they continue to be painful (11). Therefore, the ulcer is a sign of improvement while the pain may remain the same and sometimes even more painful than the initial status (Figure 1). The progression from one step to another is a sign of improvement, which is reflected by the decrease in weighting factor of lesion type in PVAS.

Also, active lesions are more prone to pressure-induce intra epithelial acantholysis on non-affected skin (direct Nikolsky’s sign). As the disease improves Nikolsky’s sign is limited around the lesion (indirect Nikolsky’s sign or Asboe-Hansen sign) and as the epithelial attachments are stabilized, no Nikolsky’s sign could be provoked (12).

In our experience, the total number of blisters on multiple sectors is more important than having the same number on one sector, especially if they are not localized on the upper body, the face, and the head. There is a large density of PV antigens in upper body areas. It seems that high systemic anti-desmoglein antibodies titration in active phase would result in widespread lesions in different anatomical regions of different PV antigen density (13). However, locally active antibodies might induce same number of bulla limited to the same sector. Koebner phenomena would be a possible example (14).

In PVAS, the average 1 cm diameter was presumed for a single blister. Large confluent and very extensive lesions are rare. They are always super infected and seen in advanced disease not responding to the treatment.

Through the process of healing, great majority of lesions in one attack progress together and go from one stage to the other. Therefore, the change of stage is a sign of improvement and the “activity index” should take it into account and decrease to show the improvement of the lesions. Finally, the total number of lesions is multiplied by the weighting factor (type of lesions).

The range goes from 0 to 18.

Here is an example for clarification: In a patient with 20 bulla on one sector of the body (on the face) without Nikolsky’s sign, the PVAS is \((1+1)\times1=2\). If no new lesions appear and two bulla with 18 crusts remain on the face, the PVDA becomes \((1+1)\times0.5=1\), showing the improvement of the disease (only the type of lesions had changed, not their number, but the disease activity is sensitive to change).

Convergent validity

Evaluation of the extent of agreement between the proposed PVAS score and experts PGA - both measure similar contest of PV activity- represents the convergent validity of the PVAS on expert opinion. Higher correlation coefficient represents stronger convergent validity.

Statistical analysis

Spearman’s correlation coefficient (Rho) represent the degree to which PVAS score is similar to (converges on) the experts opinion scores.

To detect biased experts scores the difference from PVAS was tested with paired t-test. Kolmogorov-Smirnov tested the normality of measures. SPSS version 19 was used in analysis. P-value of <0.05 was defined as significant.

Results

The mean score of the five experts was used as “expert opinion” score. The descriptive statistics are shown in Table 2.

The scatter plot showed the linear relationship between PVAS and expert opinion scores (Figure 2). The rise in expert’s opinion score was directly associated with increased PVAS score and vice versa.

Table 2. Mean scores of the five experts and PVAS. Mean score of the five experts was used as “expert opinion” score.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVAS</td>
<td>5.230</td>
<td>3.57</td>
<td>50</td>
</tr>
<tr>
<td>expert no.1</td>
<td>5.20</td>
<td>4.21</td>
<td>50</td>
</tr>
<tr>
<td>expert no.2</td>
<td>4.74</td>
<td>3.24</td>
<td>50</td>
</tr>
<tr>
<td>expert no.3</td>
<td>7.18</td>
<td>3.17</td>
<td>50</td>
</tr>
<tr>
<td>expert no.4</td>
<td>5.67</td>
<td>3.17</td>
<td>50</td>
</tr>
<tr>
<td>expert no.5</td>
<td>4.95</td>
<td>3.55</td>
<td>50</td>
</tr>
<tr>
<td>expert opinion mean</td>
<td>5.54</td>
<td>3.17</td>
<td>50</td>
</tr>
</tbody>
</table>
Table 3. Spearman correlation coefficient of PVAS on expert’s opinion was 0.751 (P-value <0.001).

<table>
<thead>
<tr>
<th></th>
<th>PVAS</th>
<th>Expert no.1</th>
<th>Expert no.2</th>
<th>Expert no.3</th>
<th>Expert no.4</th>
<th>Expert no.5</th>
<th>Expert opinion mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVAS</td>
<td>1</td>
<td>0.556*</td>
<td>0.672*</td>
<td>0.737*</td>
<td>0.752*</td>
<td>0.64*</td>
<td>0.751*</td>
</tr>
<tr>
<td>expert no.1</td>
<td>1</td>
<td>0.712*</td>
<td>0.756*</td>
<td>0.671*</td>
<td>0.758*</td>
<td>0.87*</td>
<td></td>
</tr>
<tr>
<td>expert no.2</td>
<td>1</td>
<td></td>
<td>0.791*</td>
<td>0.766*</td>
<td>0.868*</td>
<td>0.86*</td>
<td></td>
</tr>
<tr>
<td>expert no.3</td>
<td>1</td>
<td></td>
<td>0.762*</td>
<td>0.762*</td>
<td>0.91*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>expert no.4</td>
<td>1</td>
<td></td>
<td>0.748*</td>
<td>0.74*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>expert no.5</td>
<td>1</td>
<td></td>
<td>0.748*</td>
<td>0.74*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>expert opinion</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.01 level (two-tailed)

Figure 1. A PV oral lesion, one month after treatment started. Oral erosions undergo ulceration through the process of healing. It usually takes more time to be cleared.

The Spearman correlation coefficient was analyzed. The convergent validity of PVAS on expert’s opinion, the Spearman Rho was 0.751, with 95% Confidence Interval (CI) of 0.534-0.876, and P-value <0.001 (Table 3). The mean time for scoring PVAS was 3.1 minutes (±0.2).

Discussion

Evidence based practice nowadays requires the use of validated measurements. Psychometric properties of disease activity measures were assessed only in finger count articles out of thousands in some more prevalent dermatologic diseases such as atopic eczema and plaque psoriasis (15,16). Moreover, a systematic review of near a hundred clinical trials studying pemphigus, in recent 25 years, presented more than a hundred outcome measures (17). However, the paucity of validation studies in PV may be due to rarity of the disease (4).

This study presented the PVAS measurement that is used in Iran, and assessment of convergent validity showed good correlation with quantitative experts’ opinion, PGA.

Various PV outcome measures were used in previous studies to describe the disease severity or the efficacy of therapy. In 1998, Agarwal et al. (18) developed a measurement system based on the extent of cutaneous lesions similar to the psoriasis area and severity index (PASI). They benefited from the Nikolsky’s sign as a sensitive component of disease activity (18). When using the body surface area (BSA), changes more than 10% of BSA (area equal to the entire chest) would alter only one unit of the score. Therefore, the scoring system seemed not sensitive enough to changes (19).

Herbst and Bystryn proposed a measurement mainly focused on the therapeutic dose of corticosteroids and adjuvant immunosuppressant, but the dimension of single erosions changes within affected area (20). Also,
Pemphigus vulgaris activity score

the mucosal involvement was not considered in the score (20).

A number of authors applied antibody-titer levels assessed by ELISA for cutaneous (Dsg1) and oral (Dsg3) involvement as an objective measurement (21,22). However, in short follow ups, antibody levels are often not directly correlated with disease activity. Also, antibody levels may still be detectable in a clinically non-active patient (8,23). Positive direct immunofluorescent test in the normal skin is more reliable than indirect immunofluorescence test in predicting the remaining activity of the disease in patients without lesions and on minimum dose of corticosteroid (24,25).

Japanese disease severity score was based on BSA, Nikolsky’s sign, daily new lesion development, IIF or ELISA titration and percentage of oral mucosal involvement (2). Although it grossly measures the disease activity but subjective measurements of surface involvement of body and oral mucosa makes the scores less reliable. All the items had the equal power to affect the score that makes it imprecise to classify disease severity.

Two outcome measures presented more comprehensive criteria to assess PV activity:

First, the ABSIS was proposed by Pfutze et al. in 2007 (19). It assessed the extent of skin involvement based on BSA (rule of 9%), weighted by quality of lesions. The scores ranged 0-206; 150 points for skin involvement, 11 points for oral involvement, and 45 points for subjective discomfort. It was tested in 13 patients with different severities. Oral involvement was scored through Saraswat criteria for the extent of lesions in 11 areas and discomfort while eating 9 food of different consistency (26). The study provided the intra-individual difference in the disease severity after 6 months follow up. They also plotted anti Dsg1-/Dsg3-IgG autoantibodies titers. The autoantibody titers decreased by clinical improvement, but were still detectable after 6 months and thus in discordance with the observed clinical remission at that time point (19).

ABSIS calculates only the lesions of oral mucosa. While ocular, genital and anal mucosal membranes could be involved. As mentioned earlier when erosions start to heal, the lesions improve while they continue to be painful. Therefore, the ulcer which is a sign of improvement paradoxically gets high point through subjective ABSIS score.

On the other hand, the rule of 9 for extent of the lesions is both difficult to apply and less sensitive to changes (because a large BSA should be healed in order to change one unit score) (9,19). It is also not very practical, especially for patients at the beginning of their disease, as large confluent are rare, and always seen in refractory disease with super infection.

Second, PDAI was designed by international pemphigus committee, in 2006. Scoring was based on disease involvement in selected body areas, with more attention to head and neck (50 scores). It included all mucosal membrane. The PDAI score ranged from 0 to 263; 120 points for skin activity in specified areas based on number and size of lesions, 10 points for scalp activity, and 120 points for mucosal activity and 13 points representing disease damage of post inflammatory hyperpigmentation or erythema from resolving lesion (9). PDAI did not differentiate between active lesion types of new erythema, erosion or crust, which might make it less sensitive to change.

Rosenbach et al., in 2009 assessed the reliability and convergent validity of ABSIS and PDAI with PGA (9). Ten physicians assessed fifteen patients of mainly mild-to-moderate pemphigus. Physicians scored the patients using the ABSIS scale, PDAI and PGA. The majority of physicians declared that both the PDAI and ABSIS were too difficult to be incorporated into routine practice (9). PDAI had a correlation of 0.60 (0.49–0.71) while ABSIS showed lower correlation of 0.43 (0.30–0.55) with PGA (9). In comparison, PVAS that is presented in our study had a correlation of 0.751 (0.53–0.87) with PGA. The mean time for the PDAI was 4.7 min (± 0.18) and for the ABSIS was 3.9 min (± 0.18) compared to the mean time of the PVAS as 3.1 min (± 0.2). The needed time, depends on the simplicity and number of the items to be covered. However, similar to other practical works, the time will be reduced after many sessions of scoring.

In 38 years of experience in treating large number of PV patients of all severities, we found that in addition to the number and size of the lesions, the anatomical location of the lesions are important for the evaluation of the PV disease activity. The investigations of Sison-Fonacier and Bystrin (13) demonstrated that marked regional differences in the PV antigenic expression of skin would be suggestive of the distribution of lesions in this autoimmune skin disease. They detected strong expression of PV antigens in scalp, axilla, buccal mucosa, face, and neck where they are commonly involved in PV. This finding confirmed our finding that most of our PV relapse episodes occur in the upper body, the face, and the head. It seems that in mild PV, usually head and neck are involved. While accompanying lesions in lower distribution indicate that PV antigen is becoming expressed in the normally non-
expressed antigenic skin part of the body, and the disease becomes more active. Distribution of bulla on multiple sectors of the body is more important than if the same number was localized on one sector, especially on the upper body. As soon as the blister becomes crusted the disease becomes less active. We noticed that Nikolsky’s sign is not frequent in PV but is of importance when it is present around the lesions, and especially on normal skin. The spread of lesions to other mucus membranes other than oral cavity is also important. Symptoms such as pain are not good indicators of disease activity. When the oral lesions start to heal, oral erosions undergo ulceration, and the ulceration takes more time to be healed. Therefore, the pain remains for a while, even though the disease is less active and undergoes remission.

Validating a test means how well the concept or construct (pemphigus activity) can be translated into a functioning measurable reality (27). The construct validity represents the degree to which our measure can be applied closely to the concept. In this criteria-related validity, we check the performance of our measure (PVAS) against some criterion (PGA). Here, we estimated the similarity of PVAS to expert opinion, and we presented that this model could highly converge on (PVAS) against some criterion (PGA). Here, we estimated the similarity of PVAS to expert opinion, and we presented that this model could highly converge on (PVAS) against some criterion (PGA).

This article presented a pilot study to estimate the convergent validity of the criteria with five experts for 50 patients. Further comprehensive study comparing PVAS and PDAI in long term follow ups is ongoing in ABDRC. In conclusion, PVAS could appropriately reflect the expert opinion of disease activity through objective criteria. This model can be applicable for non-expert observers during scoring patients follow ups.

References


