Evaluation of Clinical Criteria for Diagnosis of Bullous Pemphigoid

Loïc Vaillant, MD; Philippe Bernard, MD, PhD; Pascal Joly, MD, PhD; Catherine Prost, MD, PhD; Bruno Labelle, MD; Christophe Bedane, MD, PhD; Brigitte Arbeille, MD; Elisabeth Thomine, MD; Philippe Bertrand, MD; Catherine Lok, MD; Jean-Claude Roujeau, MD; for the French Bullous Study Group

**Objective:** To check the potential usefulness of clinical criteria for the diagnosis of bullous pemphigoid when state-of-the-art techniques such as Western immunoblotting, immunoprecipitation, and indirect immunofluorescence on salt-split skin or direct immunoelectron microscopy are not available.

**Design:** Comparison of the clinical criteria between 2 groups (with and without bullous pemphigoid) as defined by immunoelectron microscopy used as standard criterion, in a prospective study. Multivariate logistic regression analysis was carried out by including all items that were statistically significant (at P<.05 level) in univariate analysis.

**Setting:** Five dermatology departments in teaching hospitals.

**Patients:** The 231 patients studied had subepidermal autoimmune bullous diseases with linear IgG or C3 deposits in the basement membrane zone (157 with bullous pemphigoid, 33 with cicatricial pemphigoid, 30 with epidermolysis bullosa acquisita, 5 with lupus erythematosus, and 6 others). A second set of patients was used to calculate predictive values.

**Results:** The multivariate logistic stepwise analysis resulted in a final set of predictors that included only 4 items: absence of atrophic scars, absence of head and neck involvement, absence of mucosal involvement, and age greater than 70 years. No additional variables met the .05 significance level to enter into the model. If 3 of these 4 characteristics were present, a diagnosis of bullous pemphigoid could be made with a sensitivity of 90% and a specificity of 83%; these predictive values were calculated on a sample of 70 new cases.

**Conclusions:** With an estimated incidence of bullous pemphigoid among subepidermal autoimmune bullous diseases of 80%, the presence of 3 of the 4 significant criteria allows the diagnosis of bullous pemphigoid, with a positive predictive value of 95%. Our set of clinical criteria thus allows the diagnosis of bullous pemphigoid with good validity for both clinical practice and therapeutic trials.

Arch Dermatol. 1998;134:1075-1080
Patients and Methods

All patients who had newly diagnosed subepidermal autoimmune disease were included in this prospective study. They were examined in 3 departments of dermatology from 1983 to 1989. Criteria for enrollment were age greater than 18 years, presence of bullous dermatosis, subepidermal blister in hematoxylin-eosin examination of biopsy specimens, and presence of linear IgG or C3 deposits in the BMZ of perilesional skin detected by direct immunofluorescence. Pregnant women and patients who exhibited only IgA deposits in the BMZ were excluded. Clinical findings were prospectively recorded by means of the same standardized questionnaire that had been used in previous studies. Examination by IEM was performed by a previously described technique. Two hundred thirty-one patients met the inclusion criteria and had IEM examination. Thirty-three patients seen during the period of the study in our departments with inclusion criteria were not included in the study because of the absence of IEM examination; thus, the participation rate was 88% of all eligible patients seen during the period of the study.

Clinical Criteria

All patients were seen by one of us at the time of initial examination before initiation of treatment. The standardized clinical evaluation included pruritus; vesicles or blisters; milia; atrophic scars; traumatic blisters, defined as skin detachment induced by minimal trauma, such as Nikolsky sign in normal or perilesional skin; erythematous plaques; mucosal involvement; and alopecia. Each item was evaluated immediately after enrollment and its presence or absence noted before initiation of treatment.

The localization of blisters was precisely analyzed: presence or absence on head, neck, trunk, and lower and upper limbs. These parts of the body were subdivided and presence or absence was noted in each localization (the trunk was divided into chest, abdomen, umbilicus, upper back, lower back, and buttocks).

The standardized clinical evaluation included other items, such as onset of pruritus and blisters, number of blisters, and percentage of body skin detachment, that were not selected as criteria for this study. The clinical assessment of the patients was always performed before the diagnosis was made.

Laboratory Criteria

All patients had standard laboratory tests. For evaluation we used only eosinophilia and the presence or absence of circulating anti-BMZ antibodies detected by indirect immunofluorescence.

Groups

Patients were divided into 2 groups for evaluation of the criteria: a BP group and a non-BP group. As in other studies, the criterion standard for the diagnosis of BP in this study was IEM examination. The diagnosis of BP was made by IEM when there were strictly localized thin immune deposits in the upper lamina lucida (Figure 1). When deposits were localized in the lamina densa (Figure 2) or in the sublamina densa and the anchoring fibril zone, a diagnosis of non-BP (eg, CP, EBA, or VBSLE) was made. Many patients, particularly patients with non-BP, had serological examinations such as immunoblotting and salt-split skin indirect immunofluorescence; on 135 serum samples tested by immunoblotting, 97 were positive on epidermal extract, ie, showed reactivity with BP 230 and/or BP 180 antigens. Thus, the precise diagnosis of the subepidermal autoimmune diseases was made as follows: (1) EBA: reactivity with EBA antigen (290 or 145 kd) by immunoblotting, and immune deposits localized in the sublamina densa by IEM; (2) VBSLE: same characteristics as EBA with the additional presence of systemic lupus erythematosus; and (3) CP: localization of immune deposits on the epidermal or both dermal and epidermal sides of the split, and immune deposits localized in the lamina densa overflowing into the lamina lucida.

Statistical Analysis

Data on all patients studied were coded and entered into a computer. The frequency of each item in the BP group was compared with that in the non-BP group by means of the χ² test. Odds ratios (ORs), determined by logistic regression analysis, were used to quantify the ability of each item to predict the diagnosis of BP. An OR of 1 corresponded to the absence of predictive value. An OR greater than 1 indicated positive predictive value, whereas an OR lower than 1 indicated a negative predictive value. The statistical significance of ORs was determined by the likelihood ratio test. Ninety-five percent confidence intervals for ORs were calculated. Multivariate logistic regression analysis was carried out by first including all items with univariate analysis (at P<.05 level) and backward elimination of nonsignificant (P>.05) variables (stepwise multivariate analysis). A set of diagnostic criteria was derived from this model. The sensitivity and specificity of this set of diagnostic criteria were calculated on a sample of 70 new patients (33 men and 37 women) who fulfilled the same inclusion criteria as the previous sample and who attended the same centers from 1991 to 1996.

Results

Two hundred thirty-one patients were included in the study. Mean age at diagnosis was 73.5 years. According to IEM findings, 157 patients (86 women and 71 men) were included in the BP group and 74 (32 women and
42 men) were included in the non-BP group. The mean (±SD) age was 73.5 (±16.0) years. In the non-BP group it was possible to diagnose the disease precisely in 70 of 74 patients. Thirty-three patients had CP; 24, inflammatory EBA; 8, chronic EBA; and 5, VBSLE.

The characteristics used as independent variables were age, presence of erythematous plaques, hyper-eosinophilia, anti-BMZ antibodies, absence of atrophic scar mucosal involvement, head and neck involvement, epidermal cysts, and mechanical blisters. For all these characteristics, subsequently used to determine a set of diagnostic criteria, the differences between the 2 groups were significant at \( P < .001 \), except for the criterion of anti-BMZ antibodies, which was significant at \( P = .01 \). The sensitivity and specificity of each characteristic were then determined (Table 1). We examined the ability of each item to predict the diagnosis of BP by univariate analysis (Table 1). The OR for the absence of atrophic scars was much higher than the ORs for other criteria (nearly 3 to 4 times higher).

The multivariate logistic stepwise analysis resulted in a final set of predictors that included only absence of atrophic scars, absence of head and neck involvement, absence of mucosal involvement, and age greater than 70 years. The equation of the logistic regression function was as follows: 

\[
\logit (p) = 7.06 - 2.64 \text{scar} - 1.04 \text{head} - 1.08 \text{mucosal} - 0.06 \text{age}.
\]

This stepwise multivariate analysis demonstrated that no additional variables met the .05 significance level to enter into the model (residual \( \chi^2 = 5.97, P = .42 \)), and thus no additional variables could significantly improve the ability to predict the diagnosis of BP.

A set of diagnostic criteria was derived from this model, consisting of age greater than 70 years, absence of atrophic scars (ie, depressed scars), absence of mucosal involvement (ocular, nasal, oral, anal, or genital), and absence of head and neck blisters. We then tested this set of criteria on a new group of patients to test the hypothesis that 3 of 4 criteria allow the diagnosis of BP (\( \chi^2 = 97.5, P < .001 \)). The characteristics of the validation sample population were BP (n = 52) and non-BP (EBA, 3; CP, 15). The mean age was 77.5 (±13.9) years for patients with BP and 62 ± 23 years for the non-BP group.

Table 2 shows the sensitivity, specificity, and calculated positive and negative predictive values in cases of estimated prevalence of 70% to 90% for the diagnosis of BP among the subepidermal AIBDs, as calculated on the second sample of 70 patients. According to IEM findings, these 70 patients were divided between BP (25 men and 27 women) and non-BP (EBA, 3; CP, 15). The mean age was 77.5 (±13.9) years for patients with BP and 62 ± 23 years for the non-BP group.

**COMMENT**

Subepidermal AIBDs are blistering diseases characterized by tissue-bound and circulating autoantibodies that are directed against a component of the cutaneous membrane zone; 4 of them (BP, CP, EBA, and VBLE) have linear de-
None of the clinical criteria suggested in the literature have been validated, and our study is the first to evaluate the diagnostic value of these clinical criteria. The criteria suggested by the literature had high relative values in our study, but the ORs that we calculated suggest that their respective weights are very different. The OR for the absence of scarring was nearly 3 to 5 times greater than those of other criteria. The other clinical criteria suggested in the literature were also found in our study, with nearly the same range. However, the weighting of each criterion may appear surprising, eg, epidermal plaques without scarring, which is not from herpes gestationis, which was excluded from our study.

The most important result in our study was obtained after the logistic procedure, which allowed us to retain only 4 criteria for the diagnosis of BP. These 4 criteria are different from the main criteria suggested by the literature and are not exactly the same as those suggested by the 4 higher ORs or higher relative values. The main difference from the literature is the importance of age. The most important difference from the results of ORs is that the presence of erythematous plaques, a very important criterion for the clinical description, does not increase the diagnostic value of our 4-criteria model. Indeed, the fact that no criteria other than age, absence of scarring, mucoanal involvement, or head and neck involvement increase the diagnostic accuracy of BP is the major result of our study. This set was chosen to provide acceptable specificity and good sensitivity. If the presence of 4 characteristics suggested by our model was required, very good specificity would be obtained but with poor sensitivity, and thus the requirement of the presence of the 4 characteristics would not be useful in practical routine. To estimate the predictive value of our proposed set of criteria, we calculated its positive and negative predictive values on a group of 70 new cases. On the hypothesis that our study reflects the usual features of subepidermal AIBD, this validation of the set of criteria demonstrated the good sensitivity (90%) with acceptable specificity (83%). To calculate any predictive value, the prevalence of the disease among the study population needs to be known. In a previous study, we found that the mean annual incidence of subepidermal AIBDs in 3 French regions was 10.4 per million inhabitants. The prevalence of BP among subepidermal AIBDs (herpes gestationis and linear IgA dermatosis excluded) was between 80.5% and 82.6%. Thus we can thus estimate that in nonpregnant adults with subepidermal AIBDs (ie, presence of blisters and positive direct immunofluorescence with C3 and/or IgG at the BMZ) the positive predictive value is between 93% and 98% (Table 2). Therefore, less...
than 5% of the patients who fulfilled our suggested diagnostic criteria for BP had false-positive results. This seems acceptable both in clinical practice and for future therapeutic trials. A positive predictive value greater than 95% is acceptable because it allows inclusion of a few patients with AIBDs other than BP (in particular, a very few with EBA) in such a therapeutic trial and does not rule out too many patients. On the other hand, when a patient did not fulfill our set of diagnostic criteria, the probability that the patient really had non-BP disease was 48% to 78% (Table 2). The predictive value of our set of criteria suggests that a patient with fewer than 3 of the criteria (age less than 70 years, presence of atrophic scarring, mucosal involvement, head and neck blisters) might undergo further examinations if a precise diagnosis is needed to make a medical decision. In such cases, salt-split skin immunofluorescence could be helpful.

In conclusion, our study supports the usefulness of this set of diagnostic criteria for BP for diagnosing the individual case as well as for ensuring uniformity of groups of patients for clinical and therapeutic studies. The positive predictive value of salt-split skin immunofluorescence might be as good as that of our set of diagnostic criteria, and further studies are necessary to demonstrate this hypothesis.

Accepted for publication March 24, 1998.

This study was supported in part by a grant from the French Society of Dermatology, Paris, France.

Reprints: Loïc Vaillant, MD, Hôpital Trouseau, CHU Tours, 37044 Tours Cedex, France (e-mail: vaillant@med.univ-tours.fr).

REFERENCES

23. Bhogal BS, Black MM. Diagnosis, diagnostic and research techniques. In: Woj...

The ARCHIVES is available by request to nonfederal physicians in the United States (50 states and Washington, DC) whose official American Medical Association masterfile record shows a primary specialty of dermatology in an office- or hospital-based practice as a staff physician, resident in training, or clinical fellow.

If you meet the above qualification criteria and are not currently receiving the ARCHIVES and would like to receive it each month, you must complete a free subscription request card. To receive a request card, please write Kathryn Osten, American Medical Association, Circulation Processing Department, 515 N State St, Chicago, IL 60610 (FAX 312-464-2580). A subscription request card will be sent to you in response. If you are a resident or fellow, please include verification of your training program and a complete mailing address.

©1998 American Medical Association. All rights reserved.