

Reliability of different sets of criteria in diagnosing canine atopic dermatitis applied to a population of 250 dogs seen in a veterinary teaching hospital

Thomas Brément* , Marie J. Laly†, Daniel Combarros* , Dorian Guillemaille*, Patrick J. Bourdeau*‡ and Vincent Bruet*‡

*Unit Dermatology/Parasitology/Mycology, †School of Veterinary Medicine-ONIRIS, University of Nantes, CS 40706, Route de Gachet, 44307 Nantes Cedex 03, France

‡NP3 Unit, School of Veterinary Medicine-ONIRIS, University of Nantes, CS 40706, Route de Gachet, 44307 Nantes Cedex 03, France

Correspondence: Thomas Brément, Unit Dermatology/Parasitology/Mycology, School of Veterinary Medicine-ONIRIS, University of Nantes, CS 40706, Route de Gachet, 44307 Nantes Cedex 03, France. E-mail: thomas.brement@yahoo.fr

Background – Different sets of criteria are available in veterinary dermatology for the diagnosis of canine atopic dermatitis (cAD); there are limited data to assess their reliability.

Hypothesis/Objectives – The aim of this study was to assess the discriminative and predictive ability of four sets in diagnosing cAD in a population of dogs with different skin diseases.

Animals – Two hundred and fifty dogs examined in the dermatology unit of a veterinary teaching hospital.

Methods and materials – Dogs were diagnosed with cAD, flea infestation (FI), flea bite hypersensitivity (FBH), sarcoptic mange (SM) and other skin diseases (OD). FI, FBH and SM were pooled in one group of dermatoses mimicking cAD (MD). All nonatopic dogs were grouped as a “global population” (GP). Four criteria sets were applied to all dogs. For each set of criteria, sensitivity, specificity, positive and negative predictive values (PPV, NPV), likelihood ratios (LR) and diagnostic odds ratio (DOR) were calculated.

Results – When applied to the GP, sensitivities ranged from 54 to 83%, specificities from 68 to 75% and PPVs from 35 to 43%. NPVs ranged from 87 to 94%. LRs and DORs were poor. When applied to MD, sensitivities remained unchanged, specificities and PPV were mildly higher, whereas NPVs were mildly lower and LRs and DORs were comparable.

Conclusions and clinical importance – This study showed that each set of criteria had a low diagnostic reliability when used alone. The use of more discriminant criteria integrated into a thorough clinical approach excluding MDs should be considered.

Introduction

Canine atopic dermatitis (cAD) is an allergic skin disease of major importance in veterinary practice because it affects up to 27% of dogs and is one of the primary causes of pruritus, along with flea infestation (FI) or flea bite hypersensitivity (FBH).^{1,2} This disease is defined as a genetically predisposed inflammatory and pruritic allergic skin disease with typical clinical features associated

with immunoglobulin E antibodies most commonly directed against environmental allergens.³ Its pathogenesis is complex and clinical signs can mimic other pruritic dermatosis.⁴

Based on what had been reported in human medicine,⁵ Tom Willemse proposed, in 1986, a set of criteria as a diagnostic tool to standardize the clinical diagnosis of cAD.⁶ Later, Pascal Prélard and colleagues in 1998 and then Claude Favrot and colleagues in 2010 proposed three further sets of criteria.^{7,8} One study and one case report has highlighted that some dogs with clinical AD do not fit the published criteria.^{9,10} To the best of the authors' knowledge, there is no large-scale study that has assessed the reliability of the sets of criteria in an everyday general/first opinion or specialized/referral practice.

The aim of this study was to evaluate and compare the diagnostic effectiveness of different sets of criteria applied to different canine populations (cAD and other skin diseases, both those mimicking AD and those that did not) in a dermatology clinical practice context. For this purpose, discriminative and predictive statistical measures were used.

Abbreviations: cAD, canine atopic dermatitis; DOR, diagnostic odds ratio; FBH, flea bite hypersensitivity; FI, flea infestation; IDT, intradermal testing; LR+, positive likelihood ratio; LR–, negative likelihood ratio; MD, mimicking dermatoses; NPV, negative predictive value; OD, other skin diseases; PPV, positive predictive value; SM, sarcoptic mange; Sn, sensitivity; Sp, specificity.

Accepted 31 December 2018

Conflicts of interest: No conflicts of interest have been declared.

Sources of funding: This study was self-funded.

This study was presented in part at the European Society of Veterinary Dermatology-European College of Veterinary Dermatology annual congress, 2017, Lausanne, Switzerland. *Vet Dermatol* 2017; 28: 229 (Abstract).

Methods and materials

Criteria for the diagnosis of cAD, FI, FBH and sarcoptic mange

The selection of the cases was performed over a two year period to ensure that dogs were affected by only one dermatosis (Figure 1). Dogs with no follow-up or affected with multiple dermatoses were excluded. All dogs came from the population examined in the dermatology unit of a veterinary teaching hospital. For each dog, all the data were prospectively recorded and validated.

The cAD diagnosis was based both on exclusion and inclusion criteria without using the previously published sets of criteria. Dogs had to be free of FI or clinical signs compatible with FBH or sarcoptic mange (SM). All of them (and in-contact animals) had to be treated for fleas every four weeks for at least eight weeks with no change in the pruritus. All other possible ectoparasites were excluded on the basis of adapted sampling methods (skin scrapings, hair combing, hair plucking, ear curettage and acetate tape impressions) or serological evaluation (SM) and of the absence of clinical improvement after an adequate antiparasitic treatment. Food-induced atopic dermatitis (FIAD) was explored by feeding an elimination diet for at least eight weeks (commercial hydrolysed protein diet). Finally, at least one positive response to intradermal testing (IDT) with aeroallergens (eight to 35 allergens, Destaing Laboratories; Grasse, France) and a negative reaction to IDT with a whole-body flea extract (ARTU flea antigen, Destaing Laboratories) were mandatory for inclusion. Therefore, cases of cAD were diagnosed on the basis of excluding other primary pruritic skin diseases and including persistent primary pruritus and erythema in an allergen-test positive dog. This approach was considered the "gold standard" to diagnose dogs with cAD.

Diagnosis of FBH was diagnosed on the basis of compatible skin lesions at typical body locations, by excluding other primary pruritic skin diseases, by an immediate (20 min) and/or delayed (24–72 h) positive response to flea allergen on IDT, as well as a positive

response to insecticide treatment (disappearance of the clinical signs and pruritus). Antiflea treatment was performed on the included dogs as well as every in-contact animal using products containing spinosad, macrocyclic lactones or isoxazolines.

Diagnosis of FI of dogs was by the presence of fleas, pruritus, the exclusion of other pruritic skin diseases and a negative response to an intradermal injection of flea allergen. These dogs also had a positive response to insecticide treatment (disappearance of the clinical signs and pruritus). Antiflea treatment was performed on the included dogs, as well as every in-contact animal, using products containing spinosad, macrocyclic lactones or isoxazolines.

Diagnosis of SM was according to clinical presentation, pruritus, positive skin scrapings and/or positive serological findings (Sarcoptes-Elisa 2001® dog, AFOSA GmbH Laboratories; Blankenfelde-Mahlow, Germany) and positive response to miticide treatment (disappearance of the clinical signs and pruritus and negative test results with serological evaluation when performed). Acaricidal treatment was administered for each included dog as well as every in-contact animal, using products based either on macrocyclic lactones or isoxazolines.

Dogs were assigned to four groups (Figure 1); the cAD group was dogs with AD; the MD (mimicking disease) group comprised dogs with dermatoses mimicking AD which included FI, FBH and SM; the OD (other disease) group comprised dogs with other skin diseases; and the GP ("global population") group was composed of nonatopic dogs (i.e. those with MD or OD).

Criteria for the diagnosis of cAD

There are four published reports with proposed sets of criteria to diagnose AD in dogs;^{6–8,11} the set described by Terada *et al.* was excluded because neither sensitivity nor specificity were calculated in that study.¹¹ Criteria sets were designated "Willemse", "Prélaud and others" and "Favrot and others, set 1 or set 2".^{6–8} For all dogs the four sets of criteria were applied. Dogs were considered

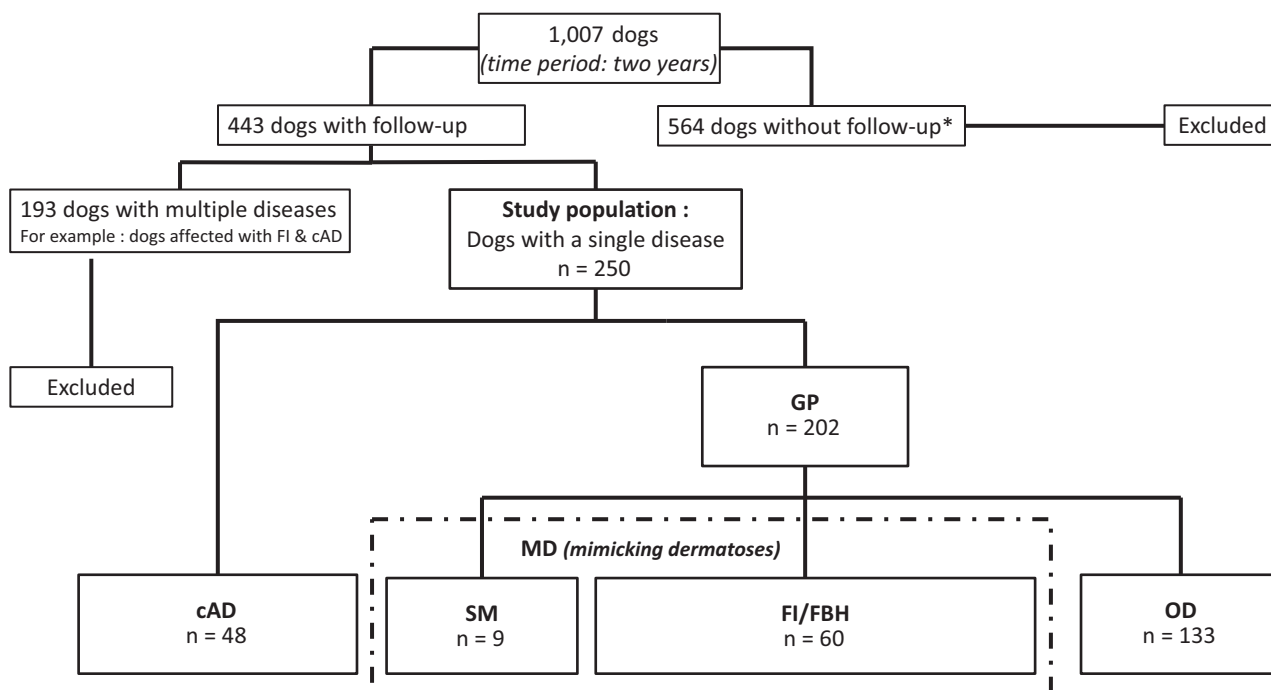


Figure 1. Inclusion/exclusion process for the study and distribution of the study population.

GP general population (nonatopic dogs), MD mimicking dermatoses, cAD canine atopic dermatitis, SM sarcoptic mange, FI/FBH flea infestation/flea bite hypersensitivity, OD other skin diseases

*The authors dermatology unit deals with primary (first opinion) cases (which have never been the subject of a previous consultation) and with referral cases. As such, the follow-up of some cases may be performed by the referring veterinarian, consequently such cases were considered as having no follow-up and were excluded from the study.

Table 1. Outline of the sets of diagnostic criteria for canine atopic dermatitis used in the study.⁶⁻⁸

Willemse ⁶	Prélaud and others ⁷	Favrot and others set 1 ⁸	Favrot and others set 2
Major criteria • Pruritus • A typical morphology and distribution ◦ Facial and/or digital involvement or ◦ Lichenification of the flexor surface of the tarsal joint and/or the extensor surface of the carpal joint • Chronic or chronically-relapsing dermatitis • An individual or family history of atopy and/or the presence of a breed predisposition Minor criteria • Onset of symptoms before the age of three years • Facial erythema and cheilitis • Bilateral conjunctivitis • A superficial staphylococcal pyoderma • Hyperhydrosis • Immediate skin test reactivity to inhalants • Elevated allergen-specific IgGd • Elevated allergen-specific IgE	Major criteria • Onset of signs between six months and three years • Glucocorticoid-responsive pruritus • Bilateral anterior interdigital erythematous pododermatitis • Erythema of the internal (concave) ear pinnae • Cheilitis Minor criteria • Breed predisposition or family history • Chronic or relapsing dermatitis (>2 years) • Dull hair • Lesions on the flexor surface of tarsal joint • Acral lick dermatitis • Hyperhydrosis • Urticaria or angioedema history • Seasonal worsening • Worsening after contact with grass • Variability of clinical signs according staying site	• Age at onset <3 years • Mostly indoor • Corticosteroid-responsive pruritus • Chronic or recurrent yeast infections • Affected front feet • Affected ear pinnae • Non-affected ear margins • Non-affected dorso-lumbar area	• Age at onset <3 years • Mostly indoor • Pruritus <i>sine materia</i> at onset • Affected front feet • Affected ear pinnae • Non-affected ear margins • Non-affected dorso-lumbar area

“positive” for the Willemse set if three or more major and three or more minor criteria were present, for the Prélaud and others set if three or more major criteria were present, and for Favrot and others sets 1 and 2 if five or more criteria were present (Table 1). The published sensitivity and specificity of the different sets are reported in Table 2.

Statistical analysis

Demographic data

Demographic data of the different subpopulations of dogs, including gender and age, were compared with previously determined data from our unit (published for FI, FBH and canine AD,¹² or unpublished for SM). Gender differences for the different subgroups of the study population were evaluated using a Pearson’s chi-square test for cAD and FI/FBH; and a Fisher’s exact test for SM (expected counts inferior to five). A Shapiro–Wilk normality test followed by a Student’s unpaired *t*-test was performed to evaluate the normal distribution of the ages and compare their mean values, respectively. For all tests, a

first type error risk of 0.05 was chosen to conclude to significant difference, associated with a 95% confidence interval. All statistical procedures were performed with the R statistical software programme.¹³

Diagnostic validity of the different sets

Various statistical measures can be used to assess the accuracy of a diagnostic test.¹⁴ These measures are often reported interchangeably without taking into account their specific ability.¹⁵ Discriminative measures are mostly used to assess the accuracy of a test and guide health-policy decisions, whereas predictive measures are more useful to predict the probability of a disease in an individual and therefore, as prognostic tools.¹⁴ To determine the overall accuracy of a diagnostic test, these statistical measures should always be analysed together and not individually.¹⁵ Measures used to assess the discriminative property of a test may include the calculation of sensitivity, specificity, likelihood ratio for positive or negative results (LR+ and LR–, respectively) and diagnostic odds ratios (DOR). To assess the predictive value

Table 2. Reported specificity, sensitivity, calculated likelihood ratios (LRs, positive and negative) and diagnostic odds ratios (DOR) of published diagnostic criteria for canine atopic dermatitis.⁶⁻⁸

	Willemse ⁶	Prélaud and others ⁷	Favrot and others ⁸ Set 1	Favrot and others ⁸ Set 2
Specificity (%)	80	81–68	79	83
Sensitivity (%)	49	79–74	85	77
LR+	2.45	4.16–2.31	4.05	4.53
LR–	0.64	0.26–0.38	0.19	0.28
DOR	3.8	16–6.1	21.3	16.2

Table 3. Demographic information about the study population

Diagnosis	Gender	Mean age (range)(years)	Main breeds
Canine atopic dermatitis (cAD)	23 m, 25 f	4.3 (one–10 years)	French bulldog West Highland white terrier Golden retriever
Flea infestation (FI) or flea bite hypersensitivity (FBH)	26 m, 34 f	5.3 (four months–13 years)	Standard poodle Labrador retriever Mixed breed
Sarcoptic mange (SM)	five m, three f	3.3 (one–nine years)	All different breeds
Other skin diseases (OD)	69 m, 64 f	4.7 (six months–14 years)	Golden retriever Labrador retriever Mixed breed

f female, m male.

of a test, measures may include the calculation of positive and negative predictive values (PPV and NPV, respectively) and diagnostic effectiveness.

Given the above, the sensitivity, specificity, PPV and NPV of each set of criteria were calculated. As PPV and NPV depend on the prevalence of the disease, they may be calculated without risk of bias only if the prevalences of the different populations of this study were similar to those of the general population seen in our dermatology unit. The prevalence of the diseases of interest (cAD, FBH, FI, SM) in our unit was known and published previously (for FI, FBH and cAD)^{16,17} or unpublished (for SM). A 95% confidence interval for a Fisher's exact test (expected counts inferior to five) was performed to compare these prevalences.

Values for LR+, LR– and DORs were calculated. LR+ shows how much more likely the positive test result is to occur in subjects with the disease compared to those without the disease. It is obtained with the formula: $LR+ = \text{sensitivity}/(1 - \text{specificity})$. The greater the LR+, the greater the indication to include the diagnosis. Diagnostic tests with a $LR+ > 10$ provide strong supportive evidence for the presence of the disease.¹⁸ LR– shows how much less likely the negative test result is to occur in a subject without the disease compared to one with the disease. It is calculated with the formula: $LR- = (1 - \text{sensitivity})/\text{specificity}$. The smaller the value of LR–, the greater the indication for ruling-out the diagnosis. Diagnostic tests with a $LR- < 0.1$ provide strong supportive evidence for the absence of the disease.¹⁸ DOR is a measure of diagnostic accuracy which estimates the overall discriminative power of a diagnostic procedure. It is the ratio of the odds of positivity in subjects with the disease to the odds in subjects without the disease, calculated with the formula: $DOR = LR+/LR-$.¹⁹ The values of a DOR range from 0 to infinity. The higher the value, the better the discriminative power of the test. A value of 1 means that the test is not able to distinguish an affected patient and an unaffected one; a value < 1 means that the test cannot be interpreted.¹⁹ LRs and DORs were calculated for the Willemse and the Favrot and others sets, using the reported values for sensitivity and specificity (Table 2).^{7,8}

Results

Population

Complete data were obtained for 250 dogs, including 48 (19%) diagnosed with cAD (no FIAD), nine (4%) with SM, 60 (24%) with FI or FBH and 133 (53%) with OD. Therefore, the GP was composed of 202 dogs. The group with OD was composed of dogs affected with neoplasia including mast cell tumours, sebaceous adenoma, epitheliotropic T-cell lymphoma, histiocytoma, lipoma ($n = 40$; 30%), fungal otitis externa ($n = 20$; 15.1%), bacterial pyoderma ($n = 19$; 14.3%), bacterial otitis externa ($n = 16$; 12.1%), ectoparasites including *Otodectes cynotis*, *Demodex* sp. and *Trombicula* sp. ($n = 11$; 8.3%), follicular dysplasia ($n = 7$; 5.2%), endocrine disorders ($n = 6$;

4.5%) and miscellaneous skin diseases including sebaceous adenitis, tick-bite associated granuloma, mucinosis and behavioural disorders ($n = 14$; 10.5%). The dogs in the OD group with bacterial/fungal otitis externa or bacterial pyoderma did not have AD, SM, flea associated disease, suspected or identified, as an underlying cause of their otitis externa and/or pyoderma.

The prevalences of cAD, FI/FBH and SM were not statistically different when compared with previously reported prevalences of the same diseases in our dermatology unit ($P = 1$).

Data regarding breeds, gender and age are reported in Table 3. No statistical difference was found for gender distribution when compared to data already known ($P = 0.25$, 0.48 and 0.40, respectively, for cAD, FI/FBH and SM). There was no statistical differences in mean age at consultation ($P = 0.65$).

Application of the different sets of criteria to the study population

Table 4 shows the number of dogs fulfilling the requirements for the different sets of criteria in dogs diagnosed with cAD, FI/FBH, SM, MD or OD.

Specificity and sensitivity

The results for sensitivity and specificity are presented in Table 5.

Positive and negative predictive values

Results of PPV and NPV for the different sets of criteria when compared to GP and MD groups are presented in Table 6.

Likelihood and diagnostic odds ratios

The results for LR+, LR– and DORs are presented in Table 5. Neither LR+ nor LR– results reached the reliability thresholds of > 10 for LR+ and < 0.1 for LR–. DORs varied between 10 and 12 for the Willemse and the Prélud and others sets, and between 4 and 6 for the two sets of criteria proposed by Favrot and others.

Discussion

The present study aimed, through different statistical measures, to assess the reliability of previously published clinical cAD diagnostic criteria. Discriminative measures depend on how the disease is defined, whereas

Table 4. Numbers of dogs positive and negative for atopic dermatitis using published diagnostic criteria compared with clinical diagnosis by group

	Willemse		Prélaud and others ⁷		Favrot and others ⁸ Set 1		Favrot and others ⁸ Set 2	
Clinical diagnosis (numbers of dog)	+	–	+	–	+	–	+	–
Canine atopic dermatitis (cAD) (n = 48)	40	8	38	10	26	22	36	12
Flea infestation/flea bite hypersensitivity (FI/FBH) (n = 60)	17	43	13	47	17	43	24	36
Sarcoptic mange (SM) (n = 9)	4	5	3	6	0	9	0	9
Mimicking dermatoses (MD) (FI/FBH + SM) (n = 69)	21	48	16	53	17	52	24	45
Other skin diseases OD (n = 133)	44	89	34	99	32	101	37	96
Study population (nonatopic dogs) (MD + OD) (n = 202)	65	137	50	152	49	153	61	141

Table 5. Discriminative statistical measures for published criteria applied to a population of nonatopic dogs and to dogs with skin diseases mimicking atopic dermatitis (cAD)

		Willemse ⁶	Prélaud and others ⁷	Favrot and others ⁸ Set 1	Favrot and others ⁸ Set 2
Study population that was not diagnosed with cAD (i.e. diagnosed with other skin diseases or dermatoses mimicking atopic dermatitis)	Sensitivity (%)	83	79	54	75
	Specificity (%)	68	75	76	70
	Positive likelihood ratio	2.59	3.2	2.23	2.48
	Negative likelihood ratio	0.25	0.28	0.61	0.36
	Diagnostic odds ratio	10.5	11.6	3.7	6.9
Mimicking diagnoses	Sensitivity (%)	83	79	54	75
	Specificity (%)	70	77	75	65
	Positive likelihood ratio	2.74	3.41	2.2	2.16
	Negative likelihood ratio	0.24	0.27	0.61	0.38
	Diagnostic odds ratio	11.4	12.6	3.6	5.6

Table 6. Predictive statistical measures: positive (PPV) and negative (NPV) predictive values of for criteria applied to a global nonatopic canine population and dogs with skin diseases mimicking canine atopic dermatitis (cAD)

		Willemse ⁶	Prélaud and others ⁷	Favrot and others ⁸ Set 1	Favrot and others ⁸ Set 2
Study population that was not diagnosed with AD (i.e. diagnosed with other skin diseases or dermatoses mimicking atopic dermatitis)	PPV (%)	38	43	35	37
	NPV (%)	94	94	87	92
Mimicking dermatoses	PPV (%)	66	70	60	60
	NPV (%)	86	84	70	79

predictive measures rely directly on the prevalence of the disease. As descriptive tools, sensitivity, specificity and LR+ and LR– were determined because they could be easily compared between the published sets of criteria and the different canine populations in the present study. As predictive tools, PPV and NPV also were calculated. To the best of the authors' knowledge, this is the first assessment of the reliability of such criteria in an everyday dermatology clinical practice with a combination of discriminative and predictive statistical tools.

Willemse used data and analysis of case reports in the human medical literature to establish a set of criteria to help standardize the diagnosis of cAD.⁶ The criteria reported by Prélaud *et al.* are deemed to be easier to apply and had an estimated 80% sensitivity and specificity.⁷ Favrot *et al.* proposed two sets of criteria with 85–77% sensitivity and 79–83% specificity (Table 2).⁸ The diagnostic accuracy of these sets has not previously been assessed, including the more recent ones, which are currently used as the reference sets. The Willemse criteria were not evaluated for sensitivity or specificity. Furthermore, the reproducibility and repeatability of the application of this set of criteria in atopic dogs according to breeds, geographical areas or type of practice were not

evaluated. The reports by Prélaud *et al.* and Favrot *et al.* included calculated specificity and sensitivity for their own sets of criteria; in addition, Favrot *et al.* also compared the specificity and sensitivity of all published sets (Table 2).^{7,8} Other accuracy measures such as PPVs and NPVs and likelihood ratio were not assessed.

The sets reported by Prélaud *et al.* and Favrot *et al.* were based on the analysis of atopic dogs recruited in multicentre studies and according to the clinician's own judgement or experience, which was considered as a gold standard to diagnose AD.^{7,8} Limitations of such an approach include introducing variability due to multiple investigators and taking cases from various geographical areas.

In our study, the four sets of criteria were tested in our population and the results compared to the reported values for sensitivity and specificity. The sensitivity of the Willemse criteria was higher, whereas the specificity was lower. The sensitivity and specificity of the Prélaud *et al.* criteria were comparable. Both sensitivity and specificity using the two criteria sets of Favrot *et al.* criteria were lower than reported. This would mean a slightly lower risk of false positive inclusion with the Willemse and Prélaud and others criteria sets, by contrast to a slightly higher

risk of false positive inclusion with both sets from Favrot and others.

The PPVs were low, suggesting that the application of the sets may lead to a significant risk of wrongly classifying dogs as having AD (57–65% risk of false inclusion). However, the NPVs were high, meaning that there is a high probability that dogs not meeting those criteria truly did not have AD (however, there was still 6–13% risk of false exclusion). PPVs were much higher when applied to a restricted population composed of dogs with dermatosis resembling cAD (MD), but NPVs were lower, showing a greater risk of false exclusion.

Theoretically, LRs directly link the pre-test and post-test probability of a disease in a specific patient.¹⁴ More precisely, this means, for example, that the smaller the LR+, the less useful the test is to confirm the suspected diagnosis. Neither LR+, LR– nor DOR depend on the disease prevalence, but rather on the clinical spectrum of disease. Consequently, these tests are reliable and their results are comparable only if the diagnostic process is the same between studies, which is the case for the present study and the ones from Willemse, Prélaud and others, and Favrot and others (i.e. there was a comprehensive investigation to establish the diagnosis of cAD). Therefore, the authors assume that the comparison of LR and DOR results with the ones calculated from previously published data was possible and relevant. The results of LR+ for the different sets show that their usefulness to include the disease is limited. Except for the set 1 of Favrot and others, the gap between the results of LR– and the reliability threshold is less important than for LR+, showing that for Willemse, Prélaud and others and set 2 of Favrot and others, the criteria can be used with a moderate efficacy to exclude the disease when a case yields negative results. When compared to the LRs calculated from previous studies, the LR results in the dogs of our study were better using the Willemse criteria (slightly increased LR+ and markedly decreased LR–), comparable with the Prélaud and others criteria, and worse with the Favrot and others sets (marked decreased LR+ and increased LR–). When LRs were assessed for MD, the results were worse for the Favrot and others criteria (lower LR+ and identical to higher LR–) and slightly better for the Willemse and Prélaud and others sets. Moreover, the DORs calculated in the present study appear overall low to moderate, confirming that it remains difficult to discriminate with certainty atopic dogs from nonatopic dogs with the strict application of these sets of criteria. More precisely, DORs for the Favrot and others criteria were lower than the ones calculated from previous studies, higher for Willemse and comparable for the Prélaud and others criteria. These discrepancies between studies may be due to higher variability in the severity and clinical presentation of our cases compared with the selected cases of cAD in previous studies.

The results of this study confirm that none of the different sets of criteria can be considered as a true gold standard to which atopic dogs should be compared. In clinical practice, a dog meeting those criteria does not necessarily have AD and on the contrary, failure to fulfil those criteria may wrongly rule out the diagnosis of AD. Indeed, breed-specific phenotypes and manifestations of cAD not

fitting the published criteria can occur and be missed or wrongly excluded by the restricted use of those criteria.^{9,10,20} In one study, $\leq 25\%$ of cases did not fit Favrot and others criteria and could not be classified as AD out of 300 dogs.¹⁰ Among them, 13% had only dorsolumbar pruritus and could have been wrongly classified as typical cases of FBH. This means that dorsolumbar pruritus or lesions do not represent an exclusion criterion for cAD and that one individual dog could have the two diseases.⁷

Therefore, it would make sense to look for other criteria, used upstream in the diagnostic process, to improve the reliability of the different sets of criteria that could be applied to selected candidates, thus giving a greater opportunity to be diagnosed as truly having AD. A few studies suggest that clinical elements such as the characteristics of pruritus, the lesional pattern and scratch reflexes could be helpful.^{20–22} Some body sites and manifestations of pruritus have been shown to be highly specific for one dermatosis.²⁰ Moreover, suggestive lesional patterns could help discriminate dermatoses that may present similarly, such as SM, AD and FBH.²¹ The pinna-pedal reflex test has been shown to be a highly specific test with a very good NPV in diagnosing SM but was not statistically relevant for the diagnosis of AD and flea associated dermatoses.^{20,22}

As precision medicine develops, other promising diagnostic approaches for identifying AD include the use of molecular techniques on serum and blood samples, evaluating selected biomarkers or establishing a molecular profile or a metabolic fingerprint of atopic dogs versus healthy subjects.^{23–26} Whether those biomarkers are truly specific and would be reliable as a diagnostic tool is still to be determined, as well as their possible use as therapeutic targets for cAD.

The main limitations of the present study are the size of the groups of dogs, especially for dermatoses other than AD, and the restrictive geographical area for inclusion. However, the latter can also be an advantage because the variability of clinical approach is less of an influence in the diagnostic procedures and the genetic variation among breeds may be moderate. The discrepancies between our results and those published previously also could be explained by different inclusion criteria of the “global population” in our study compared with previous publications. As pruritus is only one of the criteria, nonpruritic dogs fulfilling the required number of criteria could be diagnosed as cases of cAD. Thus, some dogs without pruritus were included in this study whereas in the previous studies, dogs were recruited among a population composed only of pruritic dogs. Moreover, our general population is in accordance with a reference population composed of a wide range of skin diseases. Therefore, this population, even limited, as well as the results obtained, seems to fit to the reality of clinical practice.

In the present study, an overall estimation using combined discriminative and predictive statistical measures showed that none of the four published sets of criteria appeared to reliably diagnose cAD. It seems difficult to advise their exclusive use in diagnosing this disease when applied to the general canine population, as they are associated with high risk of false positive inclusion.

On the contrary, these sets seem to be more reliable in excluding the disease in cases of negative results. Their use in a more restricted population (MD) seems to give better results in including cAD in a reliable manner but may still not be considered as sufficient. Therefore, although these criteria may be helpful in selecting a homogenous population for research, they seem to be of relatively limited interest for the clinician, compared to a classic exclusion diagnostic approach. Diagnosis of cAD should be in combination with history, clinical presentation and a thorough investigation for the exclusion of other similarly presenting skin conditions (mainly parasitic and infectious skin diseases). Ideally, more discriminant criteria should be determined in order to use such criteria as a first-line diagnostic tool in pruritic dogs.

Acknowledgements

The authors thank Stephen D. White for his constructive advice and the revision of the manuscript, and Chantal Thorin for revising the statistical analysis part of this manuscript.

References

1. Miller WH, Griffin CE, Campbell KL. *Muller and Kirk's Small Animal Dermatology*. 7th edition. Philadelphia, WB: Saunders, 2013; 363–431.
2. Hillier A, Griffin CE. The ACVD task force on canine atopic dermatitis (I): incidence and prevalence. *Vet Immunol Immunopathol* 2001; 81: 147–151.
3. Halliwell R. Revised nomenclature for veterinary allergy. *Vet Immunol Immunopathol* 2006; 114: 207–208.
4. DeBoer DJ. Introduction: canine atopic dermatitis as an evolving, multifactorial disease. In: Noli C, Foster A, Rosenkrantz W, eds. *Veterinary Allergy*. Chichester: Wiley Blackwell, 2014; 5–7.
5. Hanifin JM, Rajka G. Diagnostic features of atopic eczema. *Acta Derm Venereol (Stockh)* 1980; 92: 44–47.
6. Willemse T. Atopic skin disease: a review and a reconsideration of diagnostic criteria. *J Small Anim Pract* 1986; 27: 771–778.
7. Prélard P, Guaguère E, Alhaidari Z *et al*. Réévaluation des critères diagnostiques de la dermite atopique. *Rev Med Vet* 1998; 149: 1,057–1,064.
8. Favrot C, Steffan J, Seewald W *et al*. A prospective study of the clinical features of chronic canine atopic dermatitis and its diagnosis. *Vet Dermatol* 2010; 21: 23–31.
9. Bensignor E, Grillon M, Mueller RS. Can atopic dermatitis be diagnosed in case of atypical clinical signs? About four clinical cases *Pratique Méd Chirurgical L'animal Compagnie* 2012; 47: 1–6.
10. Bensignor E, Merven F. A better characterization of clinical signs of canine atopic dermatitis in a specialty practice: a prospective study of 300 cases. *Vet Dermatol* 2016; 27 (Suppl1): 65 (abstract).
11. Terada Y, Nagata M, Murayama N *et al*. Clinical comparison of human and canine atopic dermatitis using human diagnostic criteria (Japanese Dermatological Association, 2009): proposal of provisional diagnostic criteria for canine atopic dermatitis. *J Dermatol* 2011; 38: 784–790.
12. Bruet V, Bourdeau PJ, Roussel A *et al*. Characterization of pruritus in canine atopic dermatitis, flea bite hypersensitivity and flea infestation and its role in diagnosis. *Vet Dermatol* 2012; 23: 487–e93.
13. R Core Team. *R: A language and environment for statistical computing*, R foundation for statistical computing, Vienna, Austria, 2018. URL <https://www.R-project.org/>.
14. Simundic AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC* 2008; 19: 203–211. Available at <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4975285/pdf/ejifcc-19-203.pdf> Accessed 01/01/2019
15. Eusebi P. Diagnostic accuracy measures. *Cerebrovasc Dis* 2013; 36: 267–272.
16. Bourdeau P, Roquet F. Canine atopy: information on 601 cases. Diagnosis and associated skin diseases. In: *Proceedings of the 21st ESVD-ECVD Congress*. Lisbon, Portugal: 2006 (poster).
17. Bourdeau P, Bruet V, Travers F. Evolution of flea infestation in dogs on an eleven years period in west of France. In: *Proceedings of the World Association for the Advancement of Veterinary Parasitology Congress*. Ghent, Belgium: 2007: 317 (abstract).
18. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004; 17: 168–169.
19. Glas AS, Lijmer JG, Prins MH *et al*. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 2003; 56: 1,129–1,135.
20. Wilhem S, Kovalik M, Favrot C. Breed-associated phenotypes in canine atopic dermatitis. *Vet Dermatol* 2011; 22: 143–149.
21. Bourdeau P, Tonus A, Bruet V. Clinical pattern of sarcoptic acariasis in dogs (*Sarcoptes scabiei*) in comparison to other pruritic skin diseases. *Vet Dermatol* 2015; 26: 312 (Abstract).
22. Mueller RS, Bettenay V, Shipstone M. Value of the pinna-pedal reflex in the diagnosis of canine scabies. *Vet Rec* 2001; 148: 621–623.
23. Bruet V, Dumon H, Bourdeau P *et al*. Prospective pilot study to detect dogs with non food-induced canine atopic dermatitis using Fourier transform infrared spectroscopy. *Vet Dermatol* 2016; 27: 356–e89.
24. Almela RM, Rubio CP, Ceron JJ *et al*. Selected serum oxidative stress biomarkers in dogs with non food-induced and food-induced atopic dermatitis. *Vet Dermatol* 2018; 29: 229–e82.
25. Majewska A, Gajewska M, Dembele K *et al*. Lymphocytic, cytokine and transcriptomic profiles in peripheral blood of dogs with atopic dermatitis. *BMC Vet Res* 2016; 12: 174.
26. Asahina R, Maeda S. A review of the roles of keratinocyte-derived cytokines and chemokines in the pathogenesis of atopic dermatitis in humans and dogs. *Vet Dermatol* 2017; 28: 16–e5.

Résumé

Contexte – Différentes séries de critères sont disponibles en médecine vétérinaire pour le diagnostic de la dermatite atopique canine (cAD) ; il existe un nombre limité de données pour déterminer leur validité.

Hypothèses/Objectifs – Le but de cette étude est de déterminer la capacité discriminative et prédictive de quatre séries de diagnostic de cAD au sein d'une population de chiens atteints de différentes dermatoses.

Sujets – Deux cent cinquante chiens examinés dans l'unité de dermatologie d'un hôpital d'enseignement.

Matériel et méthode – Les chiens ont été diagnostiqués avec cAD, une infestation par les puces (FI), une hypersensibilité aux piqûres de puces (FBH), une gale sarcoptique (SM) ou d'autres dermatoses (OD). FI, FBH et SM étaient réunis en un groupe de dermatoses mimant la cAD (MD). Tous les chiens non atopiques étaient regroupés en tant que « population globale » (GP). Quatre séries de critères ont été appliquées à tous les chiens. Pour chaque série de critères, sensibilité, spécificité, valeurs prédictive négative et positive (PPV, NPV), ratio de probabilité (LR) et rapport de cotes (DOR) ont été calculés.

Résultats – Quand appliqué à GP, les sensibilités variaient de 54 à 83%, les spécificités de 68 à 75% et les PPVs de 35 à 43%. Les NPVs allaient de 87 à 94%. Les LR et DORs étaient faibles. Quand appliqué à MD, les sensibilités restaient identiques, les spécificités et PPV étaient modérément élevées tandis que NPVs étaient modérément faibles et LR et DORs étaient semblables.

Conclusions et importance clinique – Cette étude montre que chaque série de critères a une faible valeur diagnostique lorsqu'utilisée seule. L'utilisation de critères intégrés plus discriminants dans une approche clinique globale excluant les MDs doit être considérée.

Resumen

Introducción – diferentes conjuntos de criterios están disponibles en dermatología veterinaria para el diagnóstico de la dermatitis atópica canina (cAD); existe información limitada para evaluar su fiabilidad.

Hipótesis/objetivos – el objetivo de este estudio fue evaluar la capacidad discriminativa y predictiva de cuatro series de criterios en el diagnóstico de la cAD en una población de perros con diferentes enfermedades de la piel.

Animales – Doscientos cincuenta perros examinados en la unidad de dermatología de un hospital de enseñanza veterinaria.

Métodos y materiales – los perros fueron diagnosticados con cAD, infestación por pulgas (FI), hipersensibilidad a la picadura de pulgas (FBH), sarna sarcóptica (SM) y otras enfermedades de la piel (OD). FI, FBH y SM se agruparon en un grupo de dermatosis que simulaban la cAD (MD). Todos los perros no atópicos se agruparon como una "población global" (GP). Se aplicaron cuatro conjuntos de criterios a todos los perros. Para cada conjunto de criterios, se calcularon la sensibilidad, la especificidad, los valores predictivos positivos y negativos (PPV, NPV), los cocientes de probabilidad (LR) y las relaciones de probabilidad diagnóstica (DOR).

Resultados – cuando se aplicaron al GP, las sensibilidades variaron de 54 a 83%, las especificidades de 68 a 75% y las PPVs de 35 a 43%. Los NPVs variaron de 87 a 94%. LR y DORs fueron bajos. Cuando se aplicaron a la DM, las sensibilidades se mantuvieron sin cambios, las especificidades y el PPV fueron levemente más altos, mientras que los NPVs fueron ligeramente más bajos y los LR y DOR fueron comparables.

Conclusiones e importancia clínica – este estudio mostró que cada conjunto de criterios tenía una fiabilidad de diagnóstico baja cuando se utilizaba por sí solo. Se debe considerar el uso de criterios más discriminatorios integrados en un enfoque clínico completo que excluya las dermatosis similares a dermatitis atópica.

Zusammenfassung

Hintergrund – Es sind verschiedene Zusammensetzungen an Kriterien in der Veterinärdermatologie zur Diagnose einer caninen atopischen Dermatitis (cAD) verfügbar; allerdings gibt es nur limitierte Daten über ihre Zuverlässigkeit.

Hypothese/Ziele – Das Ziel dieser Studie war es, die Unterscheidungs- und die Prognosefähigkeit von vier Sets an Kriterien bei der Diagnose der cAD in einer Population von Hunden mit verschiedenen Hauterkrankungen zu untersuchen.

Tiere – Zweihundertfünfzig Hunde, die in der Abteilung für Veterinärdermatologie eines Tierspitals untersucht worden waren.

Methoden und Material – Die Hunden waren mit cAD, Flohbefall (FI), Flohstichhypersensibilität (FBH), Sarkoptesmilbe (SM) oder anderen Hauterkrankungen (OD) diagnostiziert worden. FI, FBH und SM wurden in einer Gruppe von Dermatosen zusammengefasst, die cAD (MD) vortäuschten. Alle nicht atopischen Hunde wurden als „Globale Population“ (GP) gruppiert. Vier Sets an Kriterien wurden auf alle Hunde angewendet. Es wurde für jedes Set an Kriterien die Sensibilität, die Spezifität, positive und negative Vorhersagewerte (PPV, NPV), der Likelihoodquotient (LR) und das diagnostische Quotenverhältnis (DOR) kalkuliert.

Ergebnisse – Bei Anwendung an der GP, reichten die Sensibilitäten von 54 bis 83%, die Spezifitäten von 68 bis 75% und die PPVs von 35 bis 43%. Die NPVs reichten von 87 bis 94%. Die LR und DORs waren niedrig. Bei Anwendung an den MD blieben die Sensibilitäten unverändert, die Spezifitäten und PPVs waren geringfügig höher, während die NPVs geringfügig geringer waren und die LR und DORs vergleichbar blieben.

Schlussfolgerungen und klinische Bedeutung – Diese Studie zeigte, dass jedes Set an Kriterien eine niedrige diagnostische Verlässlichkeit aufwies, wenn es alleine angewendet wurde. Die Verwendung von besser diskriminierenden Kriterien, die in einer gründlichen klinischen Herangehensweise integriert waren, wobei MDs ausgeschlossen werden, sollte in Betracht gezogen werden.

要約

背景 – 犬アトピー性皮膚炎(cAD)の断には、獣医皮膚科学においてさまざまな基準が利用可能である。しかし、信頼性を評価するデータは限られている。

仮説/目的 – 本研究の目的は、異なる皮膚疾患の犬集団における4つのcAD診断セットの識別能力および予測能力を評価することであった。

被験動物 – 獣医教育病院の皮膚科で診察を受けた250頭の犬。

材料と方法 – 犬を、cAD、ノミの侵入(FI)、ノミ刺咬性過敏症(FBH)、疥癬(SM)およびその他の皮膚疾患(OD)と診断した。FI、FBHおよびSMを、模倣cAD(MD)として1つの皮膚疾患群にプールした。すべての非アトピー犬を、「包括的な個体数」(GP)として分類した。4つの基準セットを全ての犬に適用した。それぞれの診断基準ごとに、感度、特異度、陽性および陰性予測値(PPV、NPV)、尤度比(LR)、および診断オッズ比(DOR)を算出した。

結果 – GPに適用した場合、感度は54~83%、特異度は68~75%、PPVは35~43%の範囲であった。NPVは87~94%の範囲であった。LRとDORは不良であった。MDに適用した場合、感度は変化せず、特異度とPPVはやや高く、NPVはやや低く、LRとDORは同程度であった。

結論と臨床的重要性 – 本研究では、各診断基準セットを単独で使用した場合、診断信頼性が低いことを示した。MDを除外した徹底的な臨床アプローチに統合した、より識別能力の高い基準の使用を検討すべきであった。

摘要

背景 – 兽医皮肤病学对犬异位性皮炎(cAD)的诊断标准有多种;评估其可靠性的资料有限。

假设/目标 – 本研究目的是评估四套诊断标准从不同皮肤病犬群中鉴别和预测出cAD的能力。

动物 – 250只在兽医教学医院皮肤科接受检查的犬。

方法和材料 – 诊断患有cAD、跳蚤感染(FI)、跳蚤叮咬过敏症(FBH)、疥螨(SM)和其他皮肤病(OD)的犬。将FI、FBH和SM合并于cAD(MD)相似皮肤病组,所有非异位性犬被归为“总群体”(GP)。四套标准应用于所有犬。对于每套标准,计算其敏感性、特异性、阳性和阴性预测值(PPV,NPV)、似然比(LR)和诊断优势比(DOR)。

结果 – 当应用于GP时,敏感性范围为54%-83%,异位性为68%-75%,PPV为35%-43%,NPV的范围为87%-94%,LR和DOR很低;当应用于MD时,敏感性保持不变,异位性和PPV略微升高,而NPV略低,LR和DOR接近。

结论和临床价值 – 研究表明,单独使用时,每套标准的诊断可靠性都很低。为了排除MD,应考虑在全面的临床方法中融入更多的鉴别标准。

Resumo

Contexto – Diferentes conjuntos de critérios estão disponíveis na dermatologia veterinária para o diagnóstico da dermatite atópica canina (DAC). Entretanto, são limitados os dados avaliando a sua confiabilidade.

Hipótese/Objetivos – O objetivo deste estudo foi avaliar as habilidades discriminativas e preditivas de quatro conjuntos de critérios para o diagnóstico da DAC em uma população de cães com dermatopatias diversas.

Animais – Duzentos e cinquenta cães atendidos no serviço de dermatologia veterinária de um hospital escola.

Métodos e materiais – Os cães foram diagnosticados com DAC, pulicose, dermatite alérgica à saliva de pulga (DAPP), escabiose e outras dermatopatias (OD). Pulicose, DAPP e escabiose foram agrupadas como dermatopatias que mimetizam a DAC (DM). Todos os cães não atópicos foram agrupados como “população global” (PG). Os quatro conjuntos de critérios foram aplicados a todos os cães. Para cada conjunto de critérios, foram calculados sensibilidade, especificidade, valores preditivos positivos e negativos (VPP, VPN), razões de verossimilhança (RV) e odds ratio diagnóstico (ORD).

Resultados – Quando aplicado à PG, as sensibilidades variaram entre 54 e 83%, especificidades entre 68 e 75% e o VPP entre 35 e 43%. Os VPNs variaram entre 87 e 94%. As RVs e ORDs foram baixas. Quando aplicados aos cães com DM, as sensibilidades se mantiveram intactas, já as especificidades e os VPPs foram discretamente mais altos, enquanto os VPNs foram discretamente mais baixos e a RV e ORDs foram comparáveis.

Conclusões e importância clínica – Este estudo demonstrou que todos os conjunto de critérios apresentaram baixa confiabilidade quando utilizados sozinhos. A utilização de critérios mais rigorosos integrado a uma conduta clínica excluindo as DM deve ser considerada.