

Canine Atopic Dermatitis Diagnostic Criteria: Evaluation of Four Sets of Published Criteria Among Veterinary Students

Lucille Le Roy ■ Sophie Le Poder ■ Loïc Desquilbet ■ Sebastien Perrot ■ Paola Cavana ■ Geneviève Marignac

ABSTRACT

Canine atopic dermatitis (cAD) is a major teaching point as its diagnosis and treatment are difficult. During 11 weeks, 140 dogs and students (third, fourth and fifth years) were recruited and paired. One of the four lists of diagnostic criteria was randomly attributed to each student. Concordance results, calculated with Cohen's kappa, ranged from slight ($\kappa = 0.07$) to moderate ($\kappa = 0.53$). Favrot's diagnostic criteria received the best results. It has been observed that results are improved with clinical experience. We observed that students often forgot that Favrot's criteria apply only to pruritic dogs and that the fulfillment of the criteria allows only a suspicion, not a diagnosis, of cAD. Primary pruritus and corticosteroid-responsive pruritus were often misunderstood.

Key words: dermatology, atopic dermatitis, diagnostic criteria, veterinary education, educational, carnivorous, dog

INTRODUCTION

Canine atopic dermatitis (cAD) is a major teaching objective in Veterinary Dermatology due to its frequency (up to 10% in general practice)¹ and its chronicity. CAD is defined as a "genetically predisposed inflammatory and pruritic allergic skin disease with characteristic clinical features associated with IgE antibodies most commonly directed against environmental allergens."^{2(p.207)} It is differentiated from atopic-like dermatitis, which has the same definition except that no specific IgE antibodies can be documented. However, in light of the importance of barrier dysfunction deciphered recently in both human and canine atopic dermatitis, high IgE levels are now considered to be a consequence and not a cause of atopic dermatitis. In addition, environmental conditions, including food, seem also to play a role in the development of cAD.³⁻⁵ Therefore, in this study the term cAD will refer to cAD *sensu stricto* and to atopic-like dermatitis.

As no reliable laboratory test is available, the diagnosis of atopic dermatitis is based on ruling out other similar diseases. This point is clearly difficult for students. Lists of reliable discriminators for atopic dermatitis have been developed in human medicine.³ Following their example, several lists of diagnostic criteria have been established in veterinary medicine to help in the diagnostic process.⁶⁻⁹ For Willemse's list⁹ (Table 1), the fulfillment of at least three major and three minor criteria is necessary to suspect cAD. In Prélard's list⁶ (Table 1), at least three major

criteria are necessary. In both Favrot's lists⁷ (Table 2), five or more criteria are necessary.

AIM OF THE STUDY

The main objective of the present study was to evaluate which of the published lists of diagnostic criteria was most likely to help veterinary students to include cAD in their differential diagnosis when pertinent and to avoid dismissing a likely case. We also wanted to identify specific difficulties students faced while learning how to diagnose cAD.

MATERIALS AND METHODS

Study Design and Population

The present study was conducted in a university veterinary hospital for 11 weeks during the fall semester. During this semester, third-year students have dermatology lectures and tutorials. Their clinical teaching is restricted to vaccination. Fourth- and fifth-year students were recruited during their dermatology rotation and third-year students were recruited during their preventive medicine rotation. Each student was paired with one dog presented at the dermatology consultation. Most dogs included were presented for the first time in our hospital. When the dog was presented for a follow-up consultation, only a third-year student with no access to its medical record could be paired with it. Each dog and each student were only included

Table 1: Lists of criteria published by Willemse T. (1986)⁹ and Prélaud P. et al. (1998)⁶

	Willemse T. (1986) ⁹ criteria	Prélaud P. et al. (1998) ⁶ criteria
Major features	• Pruritus • Typical morphology and distribution: – Lichenification of the flexor surface of the tarsal joint and/or the extensor surface of the carpal joint; – Facial and/or digital involvement. • Chronic or chronically relapsing dermatitis • Individual or family history of atopy and/or the presence of a breed disposition	• First signs between 6 months and 3 years of age • Glucocorticoid-responsive pruritus • Bilateral cranial erythematous pododermatitis • Erythema of the pinnae • Cheilitis
Minor features	• Onset of symptoms before the age of 3 years • Facial erythema and cheilitis • Bilateral conjunctivitis • Superficial staphylococcal pyoderma • Hyperhidrosis • Immediate skin test reactivity to inhalants • Elevated allergen-specific IgG • Elevated allergen-specific IgE	• Breed disposition or family history of atopy • Chronic or chronically relapsing dermatitis for more than 2 years • Lesions on the flexor surface of the tarsal joint • Hyperhidrosis • Dull fur • Dermatitis due to licking • History of urticaria or angioedema • Symptoms worsening seasonally • Symptoms worsening after walking in grass • Symptom variation depending on location

Table 2: Sets of criteria published by Favrot et al. (2010)⁷

Set 1	Set 2
Age at onset under 3 years	Age at onset under 3 years
Mostly indoor	Mostly indoor
Corticosteroid-responsive pruritus	Pruritus sine material at onset
Chronic or recurrent yeast infection	Affected front feet
Affected front feet	Affected ear pinnae
Affected ear pinnae	Non-affected ear margins
Non-affected ear margins	Non-affected dorso-lumbar area
Non-affected dorso-lumbar area	

once. Dogs were enrolled after an oral informed consent from their owner as no intervention specific to the present study was done on the animal.

Individual information collected about the students included attendance of a cAD lecture (attendance to lectures was not mandatory), number of times they read their notes about cAD, and number of cAD clinical cases followed (at the hospital or in private practice).

Students were asked to examine the dog with the owner and fill in a questionnaire that provided one of four published lists of diagnostic criteria (Willemse's list, Prélaud's list, and Favrot's two lists). One list among the four ones was randomly assigned to each student. To avoid bias purely due to misunderstanding, short definitions were

written under the terms *cheilitis*, *hyperhidrosis*, and *primary* (in *primary pruritus*), and it was stated that *Malassezia* were yeasts. The main investigator was available to students throughout the process to clarify misunderstandings and to check each questionnaire after the student had filled it. As clinical features and breed predisposition of cAD vary between countries, we used only diagnostic criteria based on studies that included a significant number of European dogs. Therefore, Terada's list⁸ was not included in our study.

Evaluation of Student Performance and Difficulties

For each diagnostic criterion, students could answer *yes*, *no*, or *I can't answer*. If *I can't answer* was chosen, the student had to choose between one of the three following statement: (a) *both myself and the owner don't know the answer*, (b) *the owner doesn't know the answer*, or (c) *the test had never been conducted on the dog*. In addition, it was noted for future reference when the answer given by the student did not follow the instructions in the set used. For instance, one student considered that the pruritus was corticosteroid-responsive because he mixed up a steroid and another medicine. At the end, each student had to decide whether or not the dog could be considered to have cAD. The information was also recorded when the conclusion drawn by the student was not in accordance with results obtained (e.g. only four criteria fulfilled in one of Favrot's sets and the student considers the dog likely to be atopic).

As no reliable laboratory test is available to diagnose atopic dermatitis in the dog, the clinician's diagnosis

Table 3: Student learning habits and experience

	Third-year students (n = 74)	Fourth-year students (n = 46)	Fifth-year students (n = 20)
cAD lecture earlier in their program	76%*	100%	90%†
Dermatology clinical cases tutorial earlier in their program	23%	100%	100%
Attended cAD lecture	86%	54%	55%
Already followed at least one dermatology consultation at the university	1%	96%	95%
Already followed at least one dermatology consultation at a private practice	36%	30%	65%
Already followed at least one cAD case	14%	78%	100%
Had read scientific literature on cAD	5%	15%	15%

* 18 (24%) third-year students were included before the cAD lecture took place. Attendance to lecture was not compulsory at the time these students were in third year.

† 10% of students came from other schools and were fifth-year students who had chosen a semester of 100% small animal clinics

was recorded for each dog as presence, suspicion, or absence of cAD and was used as the reference. Of the participating clinicians, two were dermatology diplomates, one was a resident, one was a certificate holder, and one was studying for the certificate. All have 3 to 15 years experience in clinical teaching. Each clinician assesses students' ability to use cAD sets several times daily in the clinics. Of the sets, they mainly used Willemse, Prélud, or Favrot set 1. The clinicians' diagnoses also involved their clinical experience (history taking, clinical examination) and results of laboratory tests.

Statistical Analysis

The κ (kappa) coefficient was used to measure the agreement between students and dermatology clinicians. We interpreted κ values as follows: >0.80 almost perfect, 0.61–0.80 substantial, 0.41–0.60 moderate, 0.21–0.40 fair, and <0.21 slight. Negative values were not considered different than random. Concordance results were calculated using the final diagnosis given by the students ("crude results"). In addition, "modified concordance results" used the results that would have been obtained if all students had provided their diagnoses according to the list of diagnostic criteria they were provided with.

If 15% or more of the students selected *I can't answer*, or if their answer did not correspond to the dog's historical or clinical features, it was considered that the misunderstanding of the corresponding criterion was not related to the student but to the criterion itself.

RESULTS

Population Studied

A total of 140 dogs and students were recruited and successfully paired. There were 74 third-year students, 46 fourth-year students, and 20 fifth-year students. Dogs were aged 4 months to 15 years (average 5.5 and median 5). There were 50% male dogs. Crossbreed dogs were the most represented (16/140, 11%). Among the 46 breeds, the Labradors (10/140, 7%), French Bulldogs (6%), German Shepherds (6%), and American Staffordshire terrier (5%) were the most represented.

Student Learning Habits, Experience and Comprehension Skills

Among the third-year students, 18 (24%) were included in the study before the cAD lecture took place. Among the other students ($n = 122$), 65% had not attended the cAD lecture and 90% of the 140 students had never read any scientific literature on the subject (Table 3). However, 90% of fifth-year students ($n = 18$) had read their lesson before dermatology clinical rotation whereas only 76% of fourth-year students ($n = 35$) had done so. All fifth-year students had already seen a case of cAD whereas only 14% of the third years and 78% of the fourth years had seen a case (Table 3).

Among all students, the questions that most frequently raised difficulties (when the answer was *I can't answer* or an inappropriate answer was given) were

- Immunological testing (from 87% to 94% left unanswered in Willemse's list)
- Corticosteroid-responsive pruritus (69%, Favrot's set 1 and Prélud's list)
- Predisposed breeds (47%, Willemse's and Prélud's lists)
- Pyoderma (46%, Willemse's list)
- Chronic yeast infection (28%, Favrot's set 1)
- Symptoms variations according to season and location (18%, Prélud's list)
- History of angioedema (15%, Prélud's list)

Only the second set of Favrot's criteria did not present difficulties for the students.

Concordance Results

The concordance between the final diagnosis suggested by students and the clinician's conclusion ranged from slight (0.07) to moderate (0.53) (Table 4). The highest level of concordance (Table 4) was obtained by the first set of Favrot's criteria for all students ($\kappa = 0.52$ to 0.53) and by Favrot's set 2 for fourth- and fifth-year students ($\kappa = 0.51$). Among third-year students, Kappa values for Favrot's set 2 and for Prélud's and Willemse's criteria were all below 0.33 (fair results). Among third-year students, 5% chose a conclusion that was not in accordance with the result suggested by the diagnostic criteria, and likewise with

Table 4: Concordance results and their interpretation according to students' level and to compliance to conclusions

Criteria	Third-year students (crude concordance)	Third-year students (modified concordance)	Fourth- and fifth-year students (crude concordance)	Fourth- and fifth-year students (modified concordance)
Favrot 1	0.53 moderate	0.53 moderate	0.52 moderate	0.52 moderate
Favrot 2	0.21 fair	0.31 fair	0.51 moderate	0.20 slight
Prélaud	0.11 slight	0.23 fair	0.11 slight	−0.13 poor
Willemse	0.07 slight	0 slight	0.33 fair	0.17 slight

Modified concordance is calculated using the results that would have been obtained if all students had based their diagnoses on the decision chart associated with the list of diagnostic criteria they were tested on. For example, if a dog fulfills only four of Favrot's criteria and the student concludes that the dog is likely atopic, crude results would count this case in the cAD+ group (student's conclusion) and modified results in the cAD- group (according to Favrot's criteria).

21% of the fourth- and fifth-year student group. For third-year students, crude concordance results are equivalent (Favrot set 1, Willemse) or higher when compared to concordance results that would have been obtained if 100% of students had chosen the conclusion in accordance with their observation ("modified concordance results," Table 4). This suggests that third-year students were most often wrong to follow what could be interpreted as their intuition about the case. Except for Favrot set 1, modified concordance results are always higher than crude results together with a change of interpretation: moderate to slight (Favrot set 2), slight to poor (Prélaud), fair to slight (Willemse). Therefore, fourth- and fifth-year students seemed overall to be mostly right not to follow the interpretation chart provided with the criteria. These observations suggest that clinical experience improves diagnostic accuracy.

Students were questioned about their learning and clinical experience to assess if learning habits were associated with diagnostic accuracy. Attending a cAD lecture slightly improves diagnostic accuracy (0.57 versus of 0.51). Concordance tended to be higher for the third-year students when they had never followed a dermatology consultation with a private practitioner ($\kappa = 0.39$) compared to when they already had ($\kappa = 0.29$). However, the interpretation remained the same (fair). Such a difference was not observed among the fourth- and fifth-year students: kappa values were 0.54 and 0.53 (for those who had already followed a dermatology consultation with a private practitioner and those who had not, respectively). Other differences in learning environment could not be interpreted due to the small number of answers.

DISCUSSION

Material and Methods

In the present study, we selected lists of diagnostic criteria based on studies enrolling a significant number of Euro-

pean dogs since cAD is known to vary between areas of the world. Considering the present results, Terada's⁸ list of criteria (validated in Japan), which remains quite simple and yet includes more clinical signs than any of the other sets, probably has interesting characteristics from a pedagogical point of view. The published version of Willemse's list of criteria was used without the *Malassezia* criterion (added in 1997).¹⁰

Even though 140 dogs and students were recruited, since four sets of criteria were screened and two populations of students were considered, too little data was collected to judge if learning habits had a real impact on diagnostic accuracy.

Taking into account that no reliable laboratory test is available to diagnose atopic dermatitis in dogs, we used the clinician's diagnosis as the gold standard, as was done in the studies published by Prélaud⁶ et al. and Favrot⁷ et al. Using only dogs that respond to allergen-specific immunotherapy, as done by Willemse⁹ et al., has the advantage of including only atopic dogs but the disadvantage of excluding a population of atopic dogs that don't respond to this treatment, and those who would not be prescribed such a treatment. Moreover, it takes four to nine months before improvement is visible. The present study is merely meant to explore students' performance and difficulties, this simple but recognized gold standard was thus considered the most appropriate for that aim.

When asked if they thought the dog was suffering from atopic dermatitis or not, 17 students (12%) raised the point that the criteria allow only to suspect atopic dermatitis as the cAD diagnosis is made after ruling out every other cause of chronic pruritus. For these students, the investigator (LLR) rephrased the question: "Would you put cAD in one of the 3 major diagnoses for this dog?" This result implies that the other students (88%) did not understand fully how to use the criteria, which when fulfilled allow for a suspicion of cAD but not a diagnosis. Among

the fifth-year students, 35% raised the question, confirming that clinical experience allows for a more appropriate use of diagnostic criteria.

Results

The κ value was calculated based on the number of agreements and disagreements between two observers, the student and the clinician. When cAD was not likely to be the diagnosis, there were few (3/20, Favrot Set 2, third-year students) to no disagreements (Favrot, fourth- and fifth-year students). The main cause of poor concordance was when the suspicion was raised by the clinician and not recognized by the student (up to 9/15, Prelaud criteria, fourth- and fifth-year students). These results may explain the wide difference observed between Favrot's set 1 and set 2 for third-year students while the two lists are so much alike. Indeed, in this group, dogs were often suspected of having cAD by the clinician. These observations suggest there may be a risk for students to overlook some cases of cAD. Nevertheless, as discussed above, the observed tendency to underdiagnose cAD may also be related to the way the question was asked. Because the clinician's diagnosis was not solely based on the criteria, as opposed to the students' diagnosis, the discordance observed may also be due to the limitations of the criteria rather than an inappropriate use of the test.

To further evaluate students' performance and difficulties, we recorded two types of mistakes: those related to the use of criteria and those related to the conclusion drawn by the student. It would have been of interest to further describe the type of error by distinguishing those caused by data collection and those due to the student not following the indications even when data were correctly collected. Even though we have a population of 140 students, as we tested four sets of criteria, the groups would have been too small for analysis.

To our knowledge, this is the first comparison of the use of different diagnostic tools for the same disease among students in veterinary medicine. Studies performed during dentistry and osteopathy training compared inexperienced students to clinicians in a setting similar to ours, and in which they also had to rely entirely on history and physical examination to provide a diagnosis. In a study performed during osteopathy training,¹¹ a $\kappa > 0.4$ was considered a sufficient result. The same range of concordance results as ours (0.52 to 0.59) was obtained during dentistry training.¹² In the latter study, results show that when confronted with a new diagnostic tool, even a good practitioner often make mistakes. In the present study, among the four lists of diagnostic criteria provided, students were familiar only with Favrot's set 1. This fact may explain why Willemse's and Prelaud's results were discordant. Among fourth- and fifth-year students, 40% of them did not respect Prelaud's criteria instructions. Another study in dentistry¹³ suggests that the biggest difficulty for beginners is to differentiate slight symptoms from absence of lesions ($\kappa = 0.57$). It could be of interest to see if results are improved when only advanced stages of the disease are presented.

Taking into account the present results, and considering that students have a tendency to favor laboratory re-

sults and to overlook historical and clinical data, we have decided to use Favrot set 2 during the 2013–2014 university year. Emphasis is given to the dog population used by Favrot et al. in their study: criteria should be applied only on a similar population, namely dogs suffering from chronic or relapsing pruritus. A dog's condition may evolve and the clinician may need to be ready to reconsider a diagnosis of atopic dermatitis if clinical signs become chronic or relapsing.

Such diagnostic lists are developed to provide a minimum list of reliable discriminators for atopic dermatitis.¹⁴ Consequently, clinical signs and elements of the history are taught without mentioning the criteria. Favrot's set 2 is introduced only as a diagnostic tool to decide if a dog should be enrolled in the diagnostic process of atopic dermatitis or not. Five or more positive criteria don't give a diagnosis of cAD but suggest that it should be investigated.

The present study confirms that the sole application of criteria cannot replace nor compensate the expertise that clinicians develop over years.

ACKNOWLEDGMENT

The authors would like to thank the Preventive Medicine and Dermatology clinicians for allowing students to participate in this study, and Estelliane for proofreading. This study was part of the veterinary thesis of Dr. Lucile Le Roy (full text available on <http://theses.vet-alfort.fr>).

REFERENCES

- 1 Hill PB, Lo A, Eden CA, et al. Survey of the prevalence, diagnosis and treatment of dermatological conditions in small animals in general practice. *Vet Rec*. 2006;158(16):533–9. <http://dx.doi.org/10.1136/vr.158.16.533>. Medline:16632525
- 2 Halliwell R. Revised nomenclature for veterinary allergy. *Vet Immunol Immunopathol*. 2006;114(3–4):207–8. <http://dx.doi.org/10.1016/j.vetimm.2006.08.013>. Medline:17005257
- 3 Ka D, Marignac G, Desquilbet L, et al. Association between passive smoking and atopic dermatitis in dogs. *Food Chem Toxicol*. 2014;66:329–33. <http://dx.doi.org/10.1016/j.fct.2014.01.015>. Medline:24491262
- 4 Meury S, Molitor V, Doherr MG, et al. Role of the environment in the development of canine atopic dermatitis in Labrador and golden retrievers. *Vet Dermatol*. 2011;22(4):327–34. <http://dx.doi.org/10.1111/j.1365-3164.2010.00950.x>. Medline:21251098
- 5 Proverbio D, Perego R, Spada E, et al. Prevalence of adverse food reactions in 130 dogs in Italy with dermatological signs: a retrospective study. *J Small Anim Pract*. 2010;51(7):370–4. <http://dx.doi.org/10.1111/j.1748-5827.2010.00951.x>. Medline:20536692
- 6 Prelaud P, et al. Réévaluation des critères de diagnostic de la dermite atopique. *Rev Med Vet (Toulouse)*. 1998;149(11):1057–64.
- 7 Favrot C, Steffan J, Seewald W, et al. A prospective study on the clinical features of chronic canine atopic dermatitis and its diagnosis. *Vet Dermatol*. 2010;21(1):23–31. <http://dx.doi.org/10.1111/j.1365-3164.2009.00758.x>. Medline:20187911

- 8 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(8):784–90. <http://dx.doi.org/10.1111/j.1346-8138.2010.01136.x>. Medline:21434981
- 9 Willemse T. Atopic skin disease: a review and a reconsideration of diagnostic criteria. *J Small Anim Pract*. 1986;27(11):771–8. <http://dx.doi.org/10.1111/j.1748-5827.1986.tb02119.x>.
- 10 Reedy LM, Miller W, Willemse T. Urticaria, angioedema and atopy. In: *Allergic skin diseases of dogs and cats*. Philadelphia, USA: WB Saunders Company Ltd; 1997. p. 39–49.
- 11 Degenhardt BF, Snider KT, Snider EJ, et al. Interobserver reliability of osteopathic palpatory diagnostic tests of the lumbar spine: improvements from consensus training. *J Am Osteopath Assoc*. 2005;105(10):465–73. Medline:16314679
- 12 Zandona AG, Al-Shiha S, Eggertsson H, et al. Student versus faculty performance using a new visual criteria for the detection of caries on occlusal surfaces: an in vitro examination with histological validation. *Oper Dent*. 2009;34(5):598–604. <http://dx.doi.org/10.2341/08-082-L>. Medline:19830976
- 13 Piovesan C, Moro BL, Lara JS, et al. Laboratorial training of examiners for using a visual caries detection system in epidemiological surveys. *BMC Oral Health*. 2013;13(1):49. <http://dx.doi.org/10.1186/1472-6831-13-49>. Medline:24090355
- 14 Williams HC, Burney PG, Hay RJ, et al, and the The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. *Br J Dermatol*. 1994;131(3):383–96. <http://dx.doi.org/10.1111/j.1365-2133.1994.tb08530.x>. Medline:7918015

AUTHOR INFORMATION

Lucille Le Roy, veterinary practitioner, Ile de France. E-mail: lucilleeroy55@gmail.com. She is interested in clinical research. This work is part of Dr. Le Roy's veterinary thesis, available on www.vet-alfort.fr.

Sophie Le Poder, DVM, PhD, is Assistant Professor in virology, Université Paris-Est, Ecole Nationale Vétérinaire d'Alfort, Unité de Parasitologie-Mycologie Dermatologie, Maisons-Alfort, F-94704, France. E-mail: slepoder@vet-alfort.fr. She is interested in veterinary virology and educational methods.

Loïc Desquilbet, PhD, is an assistant professor in Clinical Epidemiology and Biostatistics, Université Paris-Est, Ecole Nationale Vétérinaire d'Alfort, UMR 1161 de Virologie INRA-ENVA-ANSES, Maisons-Alfort, F-94704, France. Dr Desquilbet is doing his research in aging in dogs, and provides assistance to clinicians in the analyses and interpretation of results from clinical studies in animals.

Sebastien Perrot, Pharm D, PhD, ERT, is associate professor in pharmacy and toxicology, Université Paris-Est, Ecole Nationale Vétérinaire d'Alfort, Unité épidémiologie des maladies animales infectieuses ENVA USC ANSES, Maisons-Alfort, F-94704, France. E-mail: sperrot@vet-alfort.fr. He is the head of Alfort clinical research for the animal.

Paola Cavana, DVM, resident in Veterinary Dermatology, Ecole Nationale Vétérinaire d'Alfort, IRCA, 7 avenue du général de Gaulle, Maisons-Alfort, F-94704, France E-mail: pcavana@vet-alfort.fr. She is interested in clinical research.

Geneviève Marignac, DVM, PhD, dip. ECVD, is an associate professor in dermatology, Ecole Nationale Vétérinaire d'Alfort, COMUE Paris-Est, F-94704 Maisons-Alfort, France. E-mail: gmarignac@vet-alfort.fr. She is a specialist in Veterinary Dermatology. She completed a PhD on vector borne bacteria and is now focusing on clinical research in veterinary Dermatology and educational methods.