

Responsiveness and validity of the SCORFAD, an extent and severity scale for feline hypersensitivity dermatitis

Jean Steffan*, Thierry Olivry†, Sophie L. Forster* and Wolfgang Seewald*

*Novartis Animal Health, Schwarzwaldallee 214, Basel, Switzerland, †Department of Clinical Sciences and Center for Comparative Medicine and Translational Research, NC State University, Raleigh, NC, USA

Correspondence: Jean Steffan, Novartis Animal Health, Schwarzwaldallee 214, Basel, Switzerland. E-mail: jean.steffan@novartis.com

Background – Hypersensitivity (allergic) dermatitis (HD) is commonly seen in cats, causing pruritus and various patterns of skin lesions, including at least one of the following: head and neck excoriations, self-induced alopecia, eosinophilic plaques and miliary dermatitis. Few studies have evaluated the efficacy of therapeutic interventions for feline HD, and although various scales have been considered, none has been formally validated for the assessment of disease severity and its response to therapy.

Objective – To design and validate a novel scale (SCORing Feline Allergic Dermatitis; SCORFAD) to assess the value of different criteria used as outcome measures for the treatment of feline HD and to set minimal thresholds for defining the clinical success of tested interventions.

Animals – One hundred client-owned cats.

Methods – The SCORFAD scale was designed to include the four most frequently identified lesion types in feline HD (eosinophilic plaque, head and neck excoriations, self-induced alopecia and miliary dermatitis) across 10 body regions. The extent and severity of each lesion type were graded prior to inclusion and after 3 and 6 weeks in a clinical study to compare the efficacy of two doses of ciclosporin with placebo.

Results – The SCORFAD scale was found to exhibit satisfactory content, construct, criterion and sensitivity to change. The percentage reduction in SCORFAD from baseline was determined to be the most valid assessment of clinical response. Inter- and intra-observer reliability was not assessed.

Conclusions and clinical importance – The SCORFAD scale is proposed for use as a validated tool for the assessment of disease severity and response to therapeutic interventions in clinical trials for feline HD.

Introduction

Valid and reliable outcome measures are a prerequisite for evidence-based practice of both human and veterinary medicine. Summarizing therapeutic interventions in systematic reviews and compiling them in meta-analyses can only be conducted when standardized and relevant outcome measures are used in clinical trials. However, variations in methodology frequently prevent the adequate pooling of results from different studies.

A recent review indicated that up to 20 different outcome measures have been used over the years to monitor treatment response in clinical trials enrolling humans with atopic dermatitis (AD).¹ In veterinary medicine, various scales to assess the severity of skin lesions of canine AD have been used over time, especially in studies published before 2002.² Since that year, the second version of the Canine Atopic Dermatitis Extent and Severity Index

(CADESI) has been employed frequently, although it has only had limited validation.³ In contrast, the third version of this scale (CADESI-03) was found to exhibit acceptable content, construct and criterion properties, as well as inter- and intra-observer reliability and sensitivity.⁴ Furthermore, cut-off limits were established to classify the severity of this disease as 'in remission', 'mild', 'moderate' or 'severe' AD.⁵ A reduction of 50% or greater of the lesion score has been proposed as the minimal improvement for considering an intervention to be successful.²

Pruritus is an integral part of AD and, as such, it is also an important symptom to assess when evaluating the response to therapeutic interventions in this disease. Different scoring schemes have been proposed to assess the pruritus of canine AD, which were based on a categorical scale⁶, visual analog scales (VAS)⁷ and activity⁸ or video monitoring.⁹ Visual analog scales have been tested in both pruritic dogs and those without skin disease, allowing a reliable discrimination between conditions regarded as pruritic or nonpruritic.⁷

Rarely, studies have evaluated the efficacy of therapeutic interventions in clinical trials enrolling cats with hypersensitivity dermatitis (HD). Several outcome measures have been used for assessing lesion severity; some of them described a composite lesion score

Accepted 1 May 2012

Sources of Funding: This study was self-funded.

Conflict of Interest: Jean Steffan, Sophie Forster and Wolfgang Seewald are full-time employees of Novartis Animal Health. Thierry Olivry has been a consultant with Novartis Animal Health for the duration of this study.

similar to CADESI,^{10–12} the frequency of remission¹³ or the improvement in clinical signs.¹⁴ None of these scales was formally validated, however. A Feline Dermatitis Extent and Severity Index (FEDESI) score similar to the CADESI was later proposed and tested in a pilot study using a small cat population.¹⁵

Feline HD normally manifests as one of multiple lesional patterns that affect either localized areas [e.g. eosinophilic plaque (EP) and head and neck excoriations (E)] or more extensive zones [e.g. self-induced alopecia (SA) and miliary dermatitis (MD)].^{16–18} These patterns bring much variability to observed lesions, in terms of both extent and presentation, and this heterogeneity precludes relevant comparisons between cats exhibiting any of the patterns using a scale like the FEDESI.

The purpose of this paper is to report the design and validation of a novel lesion extent and severity scale, the SCORing Feline Allergic Dermatitis (SCORFAD), to be used as one of the outcome measures in clinical trials testing therapeutic interventions in cats with HD. Furthermore, we aimed to set minimal thresholds for defining the clinical success of tested interventions.

Material and methods

Feline subjects

One hundred client-owned cats with a recorded history of HD, characterized by year-round pruritus and at least one of the characteristic lesions (MD, E, SA and EP), were enrolled from 24 veterinary practices or university hospitals across five countries in a randomized double-blind parallel designed study.¹⁹ In these cats, flea bite hypersensitivity had been ruled out by use of flea adjuvants and environmental treatment for more than 1 month. In areas that were not considered to be flea endemic, a flea adjuvant was not administered. Bacterial and fungal infections were eliminated via appropriate cytology, culture and treatment prior to enrolment. Six week food trials with novel ingredients or hydrolysed proteins were used to eliminate food allergies, although these could not be completed in 15 of 100 cats. All enrolled cats received either a placebo or ciclosporin administered at 2.5 or 7.0 mg/kg daily for 42 days.

Outcome measures

Rating of skin lesions (SCORFAD).

The extent and severity of each lesion type were graded independently prior to randomization on the day of inclusion (D0) and after three (D21 ± 3 days) and 6 weeks (D42 ± 3 days) using the SCORFAD scale. In this scale, lesions of the E, EP, MD and SA patterns were evaluated. The severity and the extent of each of these patterns was recorded on 10 body regions (head, neck, dorsal and lateral thorax, rump and tail, flanks, sternum and axilla, abdomen, perineum, fore- and hindlimbs and paws; Figure 1) using a numerical rating scale with five degrees of severity [none (0), very mild (1), mild (2), moderate (3) and severe (4)]. If several lesion patterns were present, the sum of each of the pattern scores was added to yield a SCORFAD value (i.e. a total lesion score) ranging from 0 to 16 (Table 1).

Rating of pruritus (investigator pruritus score and owner pruritus score).

A VAS was used by the owners (owner pruritus score; OPS) to assess pruritus severity at D0, D21 and D42 based on their cat's behaviour during the previous 3 days. The line was marked at one end with following descriptor: 'the cat was comfortable, grooming like any normal cat', and at the other end with 'the cat was uncomfortable, grooming all the time'. Additionally, clinician investigators estimated the cat's pruritus using a five-point numerical rating scale (investigator pruritus score; IPS; Table 2).

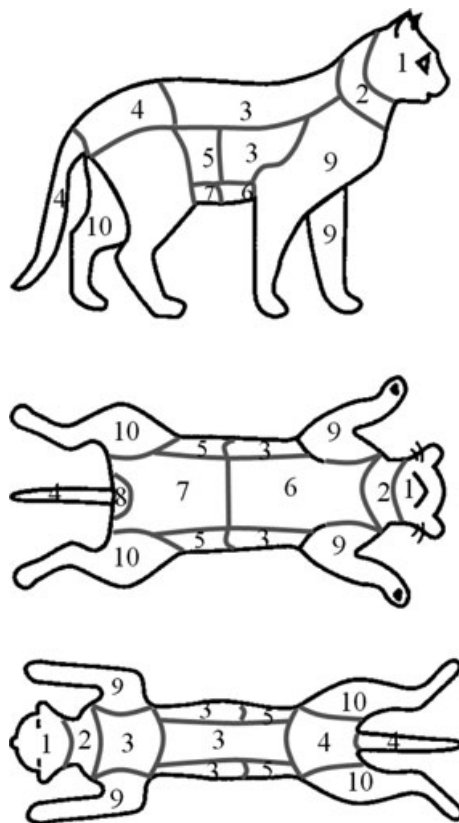


Figure 1. Body regions assessed by the SCORing Feline Allergic Dermatitis (SCORFAD) scale.

Rating of global assessment of improvement (GAI).

In addition to the parameters described above, investigators and owners determined a GAI (i.e. an overall treatment response) at study end (D42) using the numerical rating scale described in Table 3. The response to any one of the tested interventions was considered as a success when the GAI was rated 0 or 1 (excellent or good) and a failure in all other cases.

Evaluation of the SCORFAD validity

The validity of the SCORFAD was determined by assessing its content, construct and criterion properties as reported previously for the CADESI-03,⁴ as well as assessing its sensitivity to change.

Satisfying the content parameter requires that the scale appears to assess all the relevant contents or domains, based on the judgement of one or more experts. As the SCORFAD was developed by one of the authors (T.O.) with experience in scale design, this provided some relative assurance regarding the validity of the body sites, lesion types and the grading used in each section of this scale. As a result, the content characteristic of the SCORFAD was considered satisfied, and it will not be discussed further.

The construct property of a severity scale is normally assessed by determining whether the scale agrees with other related variables and measures with which, in theory, it ought to agree. In this study, the construct of the SCORFAD was studied by measuring correlations between SCORFAD and either IPS or OPS.

The criterion parameter requires a scale to correlate with some other measure of the disease in question, ideally a 'gold standard' used and accepted in the field of interest. In the absence of such a gold standard for feline HD assessment, we evaluated the criterion property by treating cats with either placebo or one of two different dosages of ciclosporin for 42 days, and by determining the correlation between the absolute change of the SCORFAD from baseline values and the percentage change of SCORFAD from baseline with the GAI of the clinical response rated by investigators.

Sensitivity to change requires a scale to demonstrate a clinically relevant change in the condition under assessment. This was carried

Table 1. The SCORing Feline Allergic Dermatitis (SCORFAD) lesion severity scale

Type of lesion	Score
Excoriations	0 = none 1 = very mild: small (≤ 1 cm long) erosion(s) in one body region 2 = mild: large (> 1 cm long) erosion(s) in one body region, or small (≤ 1 cm long) erosion(s) in more than one body region 3 = moderate: large (> 1 cm long) erosion(s) in more than one body region, or small (≤ 1 cm long) or large (> 1 cm long) ulcer(s) in one body region 4 = severe: small (≤ 1 cm long) or large (> 1 cm long) ulcer(s) in more than one body region
Miliary dermatitis	0 = none 1 = very mild: few (≤ 10) crusted papules in one body region 2 = mild: many (> 10) crusted papules in one body region 3 = moderate: few (≤ 10) crusted papules in more than one body region 4 = severe: many (> 10) crusted papules in more than one body region
Eosinophilic plaques	0 = none 1 = very mild: small (≤ 1 cm long) eroded plaque(s) in one body region 2 = mild: large (> 1 cm long) eroded plaque(s) in one body region 3 = moderate: small (≤ 1 cm long) eroded plaques in more than one body region 4 = severe: large (> 1 cm long) eroded plaques in more than one body region
Self-induced alopecia	0 = none 1 = very mild: one small (≤ 5 cm long) self-inflicted alopecic patch in one body region 2 = mild: one large (> 5 cm long) self-inflicted alopecic patch in one body region, or more than one small (≤ 5 cm long) self-inflicted alopecic patch in one body region 3 = moderate: more than one large (> 5 cm long) self-inflicted alopecic patch in one body region, or small (≤ 5 cm long) self-inflicted alopecic patches in more than one body region 4 = severe: large (> 5 cm long) self-inflicted alopecic patches in more than one body region

Table 2. Investigator pruritus scale (IPS)

0 = The cat was comfortable, grooming like any normal cat
1 = The cat was grooming, but it was tolerable and the cat remained calm
2 = The cat was grooming, but it was generally tolerable
3 = The cat was grooming quite often; the cat was uncomfortable, nervous or often agitated
4 = The cat was uncomfortable, grooming all the time

Table 3. Global assessment scale

0 = excellent (clinical signs observed during the first examination have completely disappeared)
1 = good (clear amelioration of the signs compared with initial examination)
2 = acceptable (clinical improvement compared with initial examination, but patient has responded only slightly to treatment)
3 = poor (clinical status of the cat compared with initial examination has not changed)
4 = bad (clinical status of the cat compared with initial examination has deteriorated)

out by examining baseline and final SCORFAD, as well as the absolute reduction and percentage reduction in SCORFAD following treatment with placebo, 2.5 or 7.0 mg/kg ciclosporin, to establish whether there was any difference in response according to treatment group (expectations: SCORFAD reduction placebo group < 2.5 mg/kg group < 7.0 mg/kg group).

Determination of SCORFAD outcome measures for defining treatment success

In the development of a diagnostic test, it is necessary to identify threshold values for success or failure that maximize both sensitivity (i.e. the true positive rate) and specificity (i.e. the true negative rate) of this test. An ideal test would have both high sensitivity and high specificity. In this study, sensitivity was defined as the proportion of cats with a final SCORFAD below a given cut-off value (or above a given percentage improvement from baseline SCORFAD) considered to be

a therapeutic success using the GAI score given by the investigator. In contrast, the specificity was assessed as the proportion of cats whose final SCORFAD value was above the cut-off value (or below the given improvement from baseline SCORFAD) considered to be a therapeutic failure using the GAI value attributed by the investigator.

Statistical analyses

Summary statistics were calculated for SCORFAD values pooled across all groups in the clinical study, and for the number of body regions affected in the relevant subgroups (i.e. disregarding the cats that did not have any given lesion). Pearson correlation coefficients between different lesion scores were then computed.

To determine the best threshold values that maximized both sensitivity and specificity, and therefore the accuracy of the SCORFAD scale as a measure of successful treatment, receiver operating characteristic (ROC) curves were created by plotting the sensitivity of the test against its specificity. The area under an ROC curve is an indicator of the ability of a test to discriminate success from failure. All calculations were carried out using SAS (version 9.2.2; SAS Institute Inc., Cary, NC, USA). Statistical significance was set at 5%.

Results

A total of 100 cats were included in the study. They were from a variety of breeds, predominantly domestic shorthaired (50%), and evenly distributed across treatment groups for age, sex and neuter status. A final evaluation of all outcome measures was obtained in 96 cats, which meant that when considering repeated measurements made on cats presented three times during the study, a total of 271 observations were made.

Evaluation of the SCORFAD validity

Evaluation of construct.

The OPS and IPS were significantly correlated ($P < 0.0001$), with correlation coefficients of 0.68 for ini-

Table 4. Lesion scores by lesion type

Lesion	<i>n</i>	Lesion score	Number of regions with lesions	Investigator pruritus score (IPS)	Owner pruritus score (OPS)
All lesion types	100	7.2 (3.4)	3.8 (2.3)	2.8 (0.9)	66.9 (22.6)
Excoriations	77	2.5 (0.9)	2.4 (1.6)	2.9 (0.8)	70.6 (20.9)
Miliary dermatitis	48	2.6 (1.2)	2.5 (1.7)	2.9 (0.7)	69.7 (19.1)
Eosinophilic plaques	46	2.7 (1.2)	1.9 (1.2)	3.1 (0.8)	71.2 (21.6)
Self-induced alopecia	85	3.3 (0.9)	3.2 (2.0)	2.8 (0.9)	66.5 (22.1)

Several types of lesions could be present simultaneously on a cat. All scores are given as means $\pm$ SD.

Table 5. Pruritus scores according to the number of different types of lesions present on a cat

Number of lesion types with a score of at least 2 (mild) at baseline	<i>n</i>	Investigator pruritus score (IPS)	Owner pruritus score (OPS)
		Mean (SD)	Mean (SD)
1	31	2.5 (1.0)	59.6 (25.7)
2	38	2.8 (0.8)	64.7 (22.5)
3	21	3.1 (0.7)	78.2 (17.1)
4	10	3.2 (0.6)	74.2 (12.2)
Total	100	2.8 (0.9)	66.9 (22.6)

tial and 0.82 for repeated observations. At baseline, SA had the numerically highest SCORFAD value, and affected more body regions than other lesion types (Table 4). While both IPS and OPS were of similar magnitude among the four lesion patterns (Table 4), the means of these two scores almost always increased with the increasing number of lesion types present (Table 5).

Correlations between SCORFAD and IPS or OPS were virtually identical before the therapeutic intervention ($r = 0.30$, $P = 0.0033$ and $r = 0.29$, $P = 0.0023$, respectively), thereby suggesting the existence of a weak correlation between lesion severity and pruritus. After treatment, with repeated observations being taken into

account, both correlations had increased to 0.59 ($P < 0.0001$); this reflected a parallel change in the severity of skin lesions and pruritus.

Altogether, these positive correlations satisfied the construct property of the SCORFAD.

Evaluation of criterion.

The final SCORFAD, as well as the absolute or percentage reduction in SCORFAD from baseline values were highly significantly correlated with investigator- or owner-assessed GAI ($P < 0.0001$; Table 6).

Even though the determination of pruritus by investigators or owners yielded similar results, the final OPS was better correlated than final SCORFAD or IPS with either owner or investigator GAI (data not shown). Higher correlation coefficients were obtained by measuring percentage reductions than absolute reductions from baseline, or even final SCORFAD, IPS or OPS. The highest correlation with either of the GAIs was with the percentage reduction from baseline of IPS and OPS ($r = 0.88$ for both; data not shown).

Combining measurements of lesions and pruritus in a composite score did not further improve the correlation with either owner- or investigator-assessed GAI. Nevertheless, the highest correlations were obtained by multiplying the percentage reduction of SCORFAD by either final IPS or OPS (both $r = 0.88$; Table 6).

In summary, the strong correlation between the various SCORFAD-derived outcome measures with the investigator- and owner-assessed GAI supports the validation of the criterion property of this new scale.

Evaluation of the sensitivity to change.

Full data sets were available for SCORFAD from all 100 cats enrolled in the study. The mean ($\pm$ SD) scores, absolute and percentage reduction in SCORFAD are presented in Table 7. It was clearly demonstrated that SCORFAD decreased in a stepwise fashion between cats receiving placebo, 2.5 and 7.0 mg/kg of ciclosporin. There was a significant difference in the absolute reduction in SCORFAD in favour of 7.0 mg/kg ciclosporin compared with either placebo or 2.5 mg/kg ciclosporin ($P = 0.003$ and $P = 0.0047$, respectively), whereas the difference between placebo and 2.5 mg/kg ciclosporin groups was not significant ($P = 0.44$).

Determination of SCORFAD outcome measures for defining treatment success

The threshold values of various outcome measures used to discriminate success from failure, calculated using ROC curves, are summarized in Table 8 and shown as

Table 6. Correlations of outcome measures with investigator or owner global assessment of improvement (GAI)

Variable	Correlation with investigator GAI		Correlation with owner GAI	
	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>
Final SCORFAD	98	0.61	97	0.59
Absolute reduction of SCORFAD	98	0.66	97	0.62
Percentage reduction of SCORFAD	98	0.81	97	0.75
Final SCORFAD $\times$ final IPS	98	0.67	97	0.65
Percentage reduction of SCORFAD $\times$ final IPS	98	0.88	97	0.82
Percentage reduction of SCORFAD $\times$ percentage reduction of IPS	98	0.86	97	0.81
Final SCORFAD $\times$ final OPS	96	0.65	96	0.66
Percentage reduction of SCORFAD $\times$ final OPS	96	0.88	96	0.84
Percentage reduction of SCORFAD $\times$ percentage reduction of OPS	96	0.85	96	0.82

Abbreviations: *n*, no. of feline subjects; *r*, correlation coefficient.

Table 7. The SCORFAD sensitivity to change

Response variable	Time	Placebo	Ciclosporin, 2.5 mg/kg	Ciclosporin, 7.0 mg/kg
		<i>n</i> = 35	<i>n</i> = 32	<i>n</i> = 33
SCORFAD (mean $\pm$ SD)	Day 0	7.2 $\pm$ 3.6	6.8 $\pm$ 2.8	7.7 $\pm$ 3.9
	Day 21	5.3 $\pm$ 3.3	5.0 $\pm$ 3.5	3.8 $\pm$ 2.7
	Day 42	5.6 $\pm$ 3.6	4.5 $\pm$ 3.5	2.9 $\pm$ 2.7
Absolute reduction in SCORFAD	Day 42	1.7 $\pm$ 2.8	2.3 $\pm$ 3.1	4.7 $\pm$ 4.2
Percentage reduction in SCORFAD	Day 42	20.3 $\pm$ 38.7	38.7 $\pm$ 46.8	56.1 $\pm$ 43.9

Table 8. Sensitivity and specificity based on investigator GAI

Variable	Threshold for success	Sensitivity (%)	Specificity (%)	Sensitivity $\times$ specificity (%)
Final SCORFAD*	≤ 2	74	84	62
Absolute reduction of SCORFAD†	≥ 4	71	86	61
Percentage reduction of SCORFAD†‡	40–60	86	89	77
Final IPS*	≤ 1	95	89	85
Absolute reduction of IPS†	≥ 2	95	96	71
Percentage reduction of IPS	50	83	92	76
Final OPS*‡	25–35	90	84	75
Absolute reduction of OPS†‡	20–25	78	87	68
Percentage reduction of OPS†‡	40–60	90	93	84

Success = investigator GAI 0 or 1. Failure = investigator GAI ≥ 2 . Sensitivity = proportion of cases satisfying condition among all successful cases.

Specificity = proportion of cases not satisfying condition among all unsuccessful cases.

*Considered a success if below the threshold value.

†Considered a success if above the threshold value.

‡Identical sensitivity and specificity within the range.

Figure 2. The percentage reduction in SCORFAD from baseline values was determined to be the most valid assessment of clinical response. We calculated that 86% of cats with a 40–60% SCORFAD reduction from baseline values were considered to have had a successful response to treatment of their HD (i.e. investigator-assessed GAI of 0 or 1). Additionally, 89% of cats with a SCORFAD reduction from baseline of less than 40% were considered to have a disease that had failed to respond to treatment (investigator GAI of 2–4). For both pruritus scores, a percentage improvement in the range 40–60% was also found to be the best cut-off value to discriminate success from failure (Table 8).

Discussion

In this paper, we report the design and validation of the SCORFAD, a scale developed for assessing the extent and severity of skin lesions in cats with HD. This scale was found to have acceptable content, construct and criterion, as well as being sensitive to change.

Ideally, outcome measures derived from severity scales used in clinical trials should satisfy a number of criteria to be accepted as relevant for clinical practice. They should be valid (i.e. measure what they are intended to measure), they should discriminate between situations (i.e. be reliable and sensitive to change) and be easily measurable and interpreted before being recommended.²⁰

The content validation of a scale aims at verifying that the criteria or domains used in the scale are relevant for characterizing disease severity. There is a consensus that in human AD, both lesion scores assessed by a physician

and symptoms reported by patients should be included in the core set of outcome measures in clinical trials.²¹ In the case of feline HD, four typical and frequent lesion types commonly described in the literature were selected as possible manifestations of this disease.^{11,15} In the present study, all four lesion patterns were associated with lesion and pruritus scores of similar magnitude. As expected, the extent of the lesions varied among patterns; more extensive lesions (i.e. higher numbers of regions with lesions) were found in cats with SA than in cats with EP. This observation validates *a posteriori* the design of the SCORFAD, a scale in which the extent of lesions is not considered uniformly across lesion patterns. In addition, in the present study, pruritus intensity was measured by both owner and investigator. In consensus panels, the involvement of the patient in evaluating the disease severity with a score or with a global assessment has been debated.²¹ In our study, we found an excellent correlation between pruritus scores assessed by both investigator and owner, as well as between each of these values and global assessments of HD response to treatment. As expected, in the case of a pruritic disease such as feline HD or canine AD, owners are in a better position than clinicians to assess changes in their pet's behaviour; indeed, owners have been shown to be capable of discriminating between canine pruritic and nonpruritic conditions.²²

The construct and criterion properties of a scale are normally evaluated to verify that the scale is able to characterize a disease appropriately by objective assessment of its extent and/or severity.

For construct validation, we decided to determine the correlation between two different yet possibly

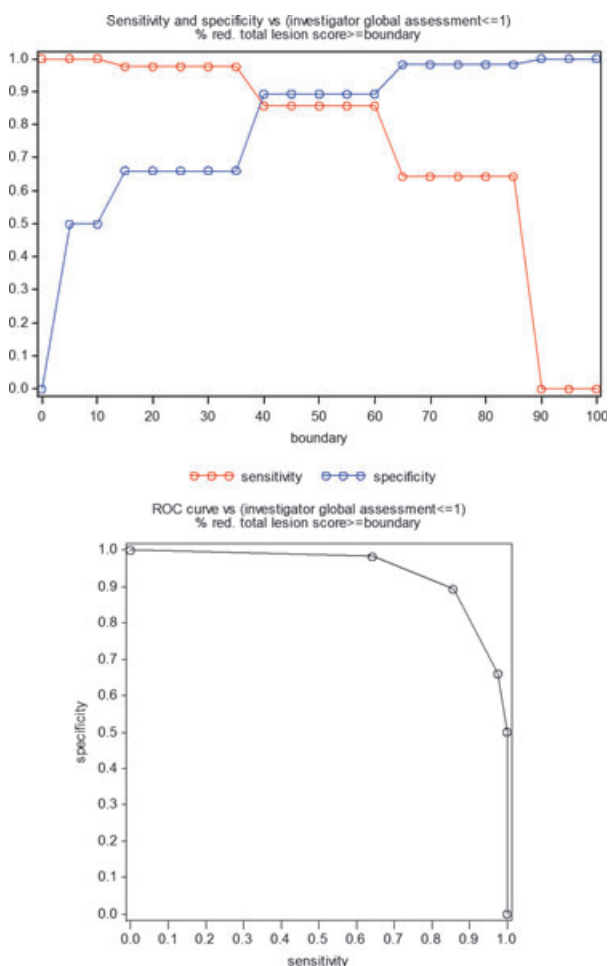


Figure 2. Sensitivity and specificity of SCORFAD versus investigator global assessment of improvement (GAI). Abbreviation: ROC, receiver operating characteristic.

related parameters: the SCORFAD and either of the two pruritus scores. Indeed, it is assumed that, in most but not all patients with allergic skin diseases, the severity of skin lesions is correlated with the severity of pruritus.²³ Our study showed that pruritus severity was not dependent on the type of lesion itself, but that it was slightly increased with higher numbers of lesion patterns present. The absence of a correlation between lesion and pruritus would have indicated that the SCORFAD was not relevant for these pruritic skin lesion patterns. At study onset, we did find a correlation of $r = 0.30$ and $r = 0.29$ between SCORFAD and IPS or OPS, respectively. At study end, with all measurements being taken into account, both pruritus scores had higher correlations of 59% with SCORFAD values. The higher correlation found after treatment might be explained by the range of responses following the therapeutic intervention in which cats received one of two different doses of ciclosporin or a placebo. However, the absence of a very high correlation between lesion and pruritus scores suggests also that these two parameters are not completely dependent on each other. As a result, they should be considered as two different aspects of allergic diseases and should be evaluated separately in clinical trials. It should also be considered that the

manifestations of pruritus in cats may be more variable and subtle than those of dogs, thereby preventing owners from making an accurate assessment of this symptom. In cats, pruritus can manifest not only as grooming but also as licking, rubbing, biting or scratching. The OPS was designed mainly for measuring the grooming behaviour, which might have underestimated pruritus intensity.

For human AD, several severity scales are considered validated (e.g. SCORAD, SASSAD and EASI), and these scales can be used as gold standards for the assessment of the criterion property.^{24–26} As an alternative method to evaluate the criterion property of the SCORFAD, we decided to measure the correlation existing between the change of this scale during treatment and the overall assessment of the clinical response to treatment made by the clinician or the owner (i.e. the GAI). Our study showed that there were highly significant correlations between various measurements of the change of SCORFAD values, the two pruritus scores or even a combination of both with the GAI. Our observations suggest that the best estimate of response to treatment, as assessed by the investigator GAI, was the percentage improvement from baseline of the SCORFAD or the OPS. The combination of the two criteria did not markedly increase this correlation, however. This observation suggests that, in clinical trials enrolling cats with HD, the evaluation of pruritus and skin lesions appears to play a similar role in the assessment of this disease, and that there is no advantage in combining these two items in a rating scale.

In this study, the SCORFAD appeared to have an appropriate sensitivity to change. This was first verified by assessing the response to various interventions, with the observation that the SCORFAD values decreased in parallel with the administration of interventions that had increasing anti-inflammatory effect (i.e. placebo and ciclosporin at 2.5 and 7.0 mg/kg).¹⁹ Furthermore, the sensitivity of this scale to change depended upon the selected outcome measures. The best sensitivity was obtained by measuring the percentage improvement from baseline of the SCORFAD. The final SCORFAD or pruritus scores or their absolute change from baseline were less sensitive than their percentage improvement from baseline values.

The reliability (i.e. intra- and interobserver repeatability) of the SCORFAD was not verified in this study. Intra-observer repeatability is normally assessed by having the investigators use the scale in the same patient at repeated intervals. The interval chosen should be sufficiently long to decrease the risk of recollection of scores given at previous time points. Moreover, one must also assume that the patient has a stable disease, and that it can be examined repeatedly; such conditions are rarely fulfilled during protocol-driven clinical trials. Assessing the interobserver repeatability is less problematic, but this parameter is also less important for clinical trials, because the same observer is normally responsible for the clinical assessment of the patient over the entire study duration. Videos or photographs could be used to overcome these limitations,⁹ but they also present inconvenience, especially when dealing with pruritus. For example, video monitoring might be able to illustrate a behaviour

observed in a given period of time that might or might not be representative of the behaviour manifested over a longer period. Indeed, a high interobserver variability is of critical importance in a multicentre study, especially if the response to the intervention is expected to be modest. If the change is low, the benefit of the intervention would be hindered by the variability of response assessment by different investigators. Alternatively, the number of subjects would have to be increased markedly for this study to have a power sufficient to detect small outcome differences.

The 'traditional' threshold used to define a successful response to therapeutic interventions in canine AD has been arbitrarily proposed to be a 50% improvement of the CADESI lesion score from baseline values. Our results show that a similar threshold can be used for assessing the response to therapeutic interventions in cats with HD. An identical threshold can also be used for the pruritus score measured by either the investigator or the owner. In contrast, a previous report provided evidence that a 50% reduction of pruritus was not sufficient to be considered satisfactory for dog owners.⁷ The difference in results between that report and the present study is probably due to a variation in methodology; in the previously reported study,⁷ the threshold was determined based on a survey in which owners were asked what improvement would be required to reach a satisfactory response. In our study, the threshold value was determined based on the response obtained after a therapeutic intervention.

In clinical trials enrolling humans with AD, a severity scale is considered adequate if each assessment does not take longer than 7 min to be completed.¹ In feline dermatology, such a duration appears conservative, particularly when considering the difficulty of restraining the subjects. In our experience acquired from an international multicentre clinical trial involving a larger number of investigators, the time spent rating lesions with the SCORFAD did not appear to be a limiting factor.

For this reason, and given that the SCORFAD satisfied the requirements reported above, we propose that the SCORFAD be considered validated for use in clinical trials testing interventions that involve enrolment of cats with HD.

Acknowledgements

The authors would like to acknowledge the contribution of all investigators who enrolled cases in the study, the owners of the cats, and the entire AtopiCat Project team at Novartis Animal Health.

References

- Schmitt J, Langan S, Williams HC. What are the best outcome measurements for atopic eczema? A systematic review. *J Allergy Clin Immunol* 2007; 120: 1389–1398.
- Olivry T, Foster AP, Mueller RS *et al.* Interventions for atopic dermatitis in dogs: a systematic review of randomized controlled trials. *Vet Dermatol* 2010; 21: 4–22.
- Olivry T, Rivierre C, Jackson HA *et al.* Cyclosporine decreases skin lesions and pruritus in dogs with atopic dermatitis: a blinded randomized prednisolone-controlled trial. *Vet Dermatol* 2002; 13: 77–87.
- Olivry T, Marsella R, Iwasaki T *et al.* Validation of CADESI-03, a severity scale for clinical trials enrolling dogs with atopic dermatitis. *Vet Dermatol* 2007; 18: 78–86.
- Olivry T, Mueller R, Nuttall T *et al.* Determination of CADESI-03 thresholds for increasing severity levels of canine atopic dermatitis. *Vet Dermatol* 2008; 19: 115–119.
- Bond R, Lloyd DH. Combined treatment with concentrated essential fatty acids and prednisolone in the management of canine atopy. *Vet Rec* 1994; 134: 30–32.
- Hill PB, Lau P, Rybníček J. Development of an owner-assessed scale to measure the severity of pruritus in dogs. *Vet Dermatol* 2007; 18: 301–308.
- Nuttall T, McEwan N. Objective measurement of pruritus in dogs: a preliminary study using activity monitors. *Vet Dermatol* 2006; 17: 348–351.
- Plant JD. Repeatability and reproducibility of numerical rating scales and visual analogue scales for canine pruritus severity scoring. *Vet Dermatol* 2007; 18: 294–300.
- Noli C, Scarpella F. Prospective open pilot study on the use of ciclosporin for feline allergic skin disease. *J Small Anim Pract* 2006; 47: 434–438.
- Wisselink MA, Willemse T. The efficacy of cyclosporine A in cats with presumed atopic dermatitis: a double blind, randomised prednisolone-controlled study. *Vet J* 2009; 180: 55–59.
- Schmidt V, Buckley LM, McEwan NA *et al.* Efficacy of a 0.0584% hydrocortisone aceponate spray in presumed feline allergic dermatitis: an open label pilot study. *Vet Dermatol* 2012; 23: 11–16, e3–e4.
- Vercelli A, Raviri G, Cornegiani L. The use of oral cyclosporin to treat feline dermatoses: a retrospective analysis of 23 cases. *Vet Dermatol* 2006; 17: 201–206.
- Nakazato A, Ishikawa H, Takizawa Y *et al.* Administration of cyclosporine to cats with allergic dermatitis: improvement of clinical signs and decrease of peripheral eosinophils in an open pilot study. *Vet Dermatol* 2006; 17: 215 (abstract).
- Nuttall TJ, Steen RV, Cawood MI *et al.* Feline Dermatitis Extent and Severity Index: a pilot study. *Vet Dermatol* 2004; 15(Suppl. 1): 36 (abstract).
- Scott DW, Miller WH, Griffin CE. Skin immune system and allergic skin diseases. In: *Muller and Kirk's Small Animal Dermatology*, 6th edn. Philadelphia: W.B. Saunders Co., 2001; 543–666.
- Favrot C, Steffan J, Seewald W *et al.* Establishment of diagnostic criteria for feline nonflea-induced hypersensitivity dermatitis. *Vet Dermatol* 2012; 23: 45–50, e11.
- Hobi S, Linek M, Marignac G *et al.* Clinical characteristics and causes of pruritus in cats: a multicentre study on feline hypersensitivity-associated dermatoses. *Vet Dermatol* 2011; 22: 406–413.
- King S, Favrot C, Messinger L *et al.* A randomised double-blinded placebo controlled study to evaluate an effective ciclosporin dose for treatment of feline hypersensitivity dermatitis. *Vet Dermatol* 2012; doi: 10.1111/j.1365-3164.2012.01086.x.
- Boers M, Brooks P, Strand CV *et al.* The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998; 25: 198–199.
- Schmitt J, Langan S, Stamm T *et al.* Core outcome domains for controlled trials and clinical recordkeeping in eczema: international multiperspective Delphi consensus process. *J Invest Dermatol* 2011; 131: 623–630.
- Rybníček J, Lau-Gillard PJ, Harvey R *et al.* Further validation of a pruritus severity scale for use in dogs. *Vet Dermatol* 2009; 20: 115–122.
- Hill P, Rybníček J, Lau-Gillard P. Correlation between pruritus score and grossly visible erythema in dogs. *Vet Dermatol* 2010; 21: 450–455.
- Barbier N, Paul C, Luger T *et al.* Validation of the Eczema Area and Severity Index for atopic dermatitis in a cohort of

- 1550 patients from the pimecrolimus cream 1% randomized controlled clinical trials programme. *Br J Dermatol* 2004; 150: 96–102.
25. Charman CR, Venn AJ, Williams HC. Reliability testing of the Six Area, Six Sign Atopic Dermatitis severity score. *Br J Dermatol* 2002; 146: 1057–1060.
26. Kunz B, Oranje AP, Labreze L *et al.* Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force on Atopic Dermatitis. *Dermatology* 1997; 195: 10–19.

Résumé

Contexte – La dermatite par hypersensibilité (allergique) (HD) est fréquemment rencontrée chez le chat et induit du prurit et des patrons lésionnels cutanés variés comprenant au moins l'un des syndromes suivants : prurit cervico-facial, alopecie extensive, plaques éosinophiliques et dermatite miliaire. Peu d'études ont évalués l'efficacité des traitements des HD félines et bien que des échelles variées aient été utilisées, aucune n'a été formellement validée pour l'évaluation de la sévérité de la maladie et sa réponse au traitement.

Objectif – Concevoir et valider une nouvelle échelle (SCORing Feline Allergic Dermatitis; SCORFAD) pour évaluer la valeur de différents critères utilisés comme mesures de suivi du traitement de HD félines et de fixer des seuils limites pour définir le succès clinique des interventions testées.

Sujets – Cent propriétaires de chats.

Méthodes – L'échelle SCORFAD a été conçue pour inclure les quatre types de lésions les plus fréquemment retrouvées dans les HD félines (plaque éosinophilique, prurit cervico-facial, alopecie extensive et dermatite miliaire) sur 10 régions corporelles. L'étendue et la sévérité de chaque type de lésion ont été notées avant l'inclusion et après 3 et 6 semaines d'une étude clinique comparant l'efficacité de deux doses de cyclosporine et d'un placebo.

Résultats – L'échelle SCORFAD s'est révélée satisfaisante sur le contenu, la construction, les critères et la sensibilité au changement. Le pourcentage de réduction du SCORFAD à partir de l'inclusion s'est révélé être la meilleure estimation de la réponse clinique. La concordance inter et intra-observateur n'a pas été évaluée.

Conclusions et importance clinique – L'échelle SCORFAD est proposée comme un outil validé pour l'évaluation de la sévérité de la maladie et la réponse au traitement dans les essais cliniques sur les HD félines.

Resumen

Introducción – la dermatitis por hipersensibilidad (alérgica) (HD) se observa con frecuencia en gatos, causando prurito y diversas presentaciones de lesiones en la piel, incluyendo por lo menos una de las siguientes: excoriaciones de la cabeza y del cuello, alopecia auto-inducida, placas eosinofílicas y dermatitis miliar. Pocos estudios han evaluado la eficacia de los tratamientos frente a la HD felina, y aunque se han considerado diversas escalas de valoración, no se ha validado ninguna formalmente para la evaluación de la severidad de la enfermedad y de su respuesta a la terapia.

Objetivo – diseñar y validar una nueva escala (valoración de la dermatitis alérgica felina; SCORFAD) para determinar el valor de diversos criterios usados como parámetros de respuesta para el tratamiento de la HD felina y para establecer los umbrales mínimos que ayuden a definir el éxito clínico de terapias en estudio.

Animales – cien gatos de propietarios particulares.

Métodos – la escala SCORFAD fue diseñada para incluir los cuatro tipos más frecuentes de lesión identificada en la HD felina (placa eosinofílica, excoriaciones de la cabeza y del cuello, alopecia auto-inducida y dermatitis miliar) distribuidas en 10 regiones del cuerpo. El grado y la severidad de cada tipo de lesión fueron calificados antes de la inclusión en el estudio y después de 3 y 6 semanas para comparar la eficacia de dos dosis de ciclosporina con placebo.

Resultados – la nueva escala SCORFAD fue satisfactoria en cuanto a contenido, construcción, criterio y la sensibilidad a cambios. La reducción del porcentaje en SCORFAD con respecto al valor basal fue el parámetro más válido de respuesta clínica. No se determinó la fiabilidad inter e intraobservador.

Conclusiones e importancia clínica – se propone la escala SCORFAD para uso como herramienta validada para la evaluación de la severidad de la enfermedad y la respuesta a las intervenciones terapéuticas en ensayos clínicos para la HD felina.

Zusammenfassung

Hintergrund – Eine Dermatitis aufgrund von Hypersensibilität (allergische Dermatitis) (HD) wird bei Katzen häufig gesehen, was Juckreiz und verschiedene Veränderungsmuster der Haut verursacht, wobei zumindest eine der folgenden Läsionen vorkommt: Exkoriationen an Hals und Kopf, selbst-verursachte Alopecie, eosinophile Plaques und miliare Dermatitis. Nur wenige Studien haben den Erfolg therapeutischer Eingriffe bei HD von Katzen untersucht und obwohl verschiedene Werteskalen überlegt worden sind, wurde bisher keine davon formal für die Erhebung der Schwere der Krankheit und einer eventuellen Verbesserung durch Behandlung validiert.

Ziel – Das Ziel dieser Studie war das Design und die Validierung einer neuen Werte-Skala (SCORing Feline allergische Dermatitis; SCORFAD), um die Aussagekraft verschiedener Kriterien, die verwendet werden,

um den Therapieerfolg bei der feline HD zu beurteilen und Mindestwerte festzulegen, um einen klinischen Erfolg der untersuchten Behandlungen zu definieren.

Tiere – Einhundert private Katzen.

Methoden – Eine SCORFAD Skala wurde entwickelt, die die vier häufigsten Typen der Hautveränderungen der feline HD (eosinophile Plaques, Exkoriationen an Hals und Kopf, selbst-verursachte Alopezie, und miliare Dermatitis) an 10 Körperstellen festlegte. Das Ausmaß und der Schweregrad einer jeden Veränderung wurden zu Beginn sowie nach 3 und 6 Wochen festgehalten, um in einer klinischen Studie die Wirksamkeit zweier Dosierungen von Ciclosporin und Placebo zu vergleichen.

Ergebnisse – Es wurde festgestellt, dass die SCORFAD Skala im Inhalt, im Aufbau, an Kriterien und an Sensibilität im Bezug auf Veränderungen ausreichend war. Der Prozentanteil der Reduzierung der SCORFAD von den Ausgangswerten war das wichtigste Utensil bei der Beurteilung einer klinischen Verbesserung. Inter- und intra-Beobachter Verlässlichkeit wurden nicht erfasst.

Schlussfolgerung und klinische Bedeutung – Es kann die Verwendung der SCORFAD Skala als validierte Methode für die Erfassung der Schwere der Erkrankung und die Reaktion derselben auf therapeutische Interventionen in klinischen Studien über feline HD empfohlen werden.

要約

背景–猫の過敏性（アレルギー性）皮膚炎（HD）は一般的な疾患で、痒痒や以下の1つ以上の皮膚病変パターンが起きる：頭部と頸部の表皮剥離、自己誘発性脱毛、好酸球性局面、粟粒性皮膚炎である。猫のHDの治療効果を評価している研究はほとんどなく、様々な基準が検討されているが、疾患の重症度と治療への反応の評価は正式に検証されていない。

目的–新規の基準(SCORing Feline Allergic Dermatitis; SCORFAD)を考案、立証するために猫のHDの治療結果判定法として用いられている様々な基準の有用性を評価し、試験し治療の臨床的な有効性を判定する最少の閾値を設定することである。

供与動物–100頭の飼い猫

方法–SCORFADは4種類の最も頻繁に見られる猫のHDの病型（好酸球性プラーク、頭部と頸部の表皮剥離、自己誘発性脱毛と粟粒性皮膚炎）を、全身の10ヶ所の部位で評価するよう計画した。試験前と試験の3週と6週後に、シクロスポリンの2種類の用量とプラセボの効果を比較するために、それぞれの病変の範囲と重症度を記録した。

結果–SCORFADは病変の変化に対して満足すべき内容、構成、基準、感受性を示した。SCORFADの初期値からの減少度は臨床的な反応に対して最も立証された評価であることがわかった。観察者間、観察者内の信頼性は評価しなかった。

結論と臨床的な重要性–SCORFADは猫のHDに対する臨床研究での疾患の重要度と介入治療に対する反応の評価に、検証されたツールとして使用できることを提案する。

摘要

背景 – 過敏（変応）性皮膚炎（HD）常見于猫，引起痒痒和多种皮肤病变模式，至少包括下列中的一种：头和颈部抓痕、自损性脱毛、嗜酸性斑块和粟粒状皮炎。评估治疗性干预猫HD有效性的研究很少，尽管认为有多种评分表，但都未被正式确认能用来评估疾病严重程度和治疗反应。

目的 – 设计并确认一个新表（猫过敏性皮炎评分；SCORFAD）评估不同标准值，作为猫HD治疗的检测结果，并设定最低阈值来定义临床治疗效果。

动物 – 一百只家养猫。

方法 – SCORFAD表的设计包括猫HD常见的四种病变类型（嗜酸性斑块、头和颈部抓痕、自损性脱毛和粟粒状皮炎），病变贯穿身体10个部位。在临床研究中每个病变类型的范围和严重程度评级，比较治疗前、3和6周后两个剂量的环孢素和安慰剂的功效。

结果 – SCORFAD表展现了多种令人满意的改变，包括内容、构想、标准和敏感性。确定SCORFAD百分比低于基础值，是临床反应的最有效评估。没有评估观察者自身和观察者之间的可靠性。

结论和临床价值 – 在临床试验中评估猫HD的严重程度和治疗反应，SCORFAD表可以做为有效工具。