Accepted Manuscript

Title: Famacha©scores should not be handled as numerical data

Author: Maurice Mahieu

PII: S0304-4017(17)30403-X
DOI: <http://dx.doi.org/10.1016/j.vetpar.2017.09.014>
Reference: VETPAR 8475

To appear in: *Veterinary Parasitology*

Received date: 21-8-2017
Revised date: 13-9-2017
Accepted date: 15-9-2017

Please cite this article as: Mahieu, Maurice, Famacha©scores should not be handled as numerical data. *Veterinary Parasitology* <http://dx.doi.org/10.1016/j.vetpar.2017.09.014>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Famacha© scores should not be handled as numerical data

Maurice MAHIEU

INRA-URZ, UR 0143, domaine Duclos, FR-97170 PETIT-BOURG

maurice.mahieu@inra.fr

Keywords:

Famacha©; score; categorical data; numerical data

The Famacha© method was developed in the late 20th century by South African parasitologists and extensionists faced with increasing difficulties in controlling *Haemonchus contortus*, a haematophagous nematode threatening small ruminant farming (Bath et al., 1996; Van Wyk and Bath, 2002). The method was originally developed in sheep, and during the early 2000s, it became popular worldwide. It was also tested successfully in goat (Kaplan et al., 2004; Mahieu et al., 2007), and has even been used for monitoring other diseases such as bovine trypanosomiasis (Grace et al., 2007).

The method relies on the visual appraisal of anaemia through examination of the colour of the ocular mucous membranes, specifically the conjunctiva on the inside of the lower eyelid. The examiner assigns each animal (often with the help of a Famacha© colour card) to one of five conventional categories or Famacha scores (FamSco), ranging from 1 — very healthy animal, no anaemia, to 5 — severe anaemia (FamSco), from 1 - very healthy animal, no anaemia - to 5 - severe anaemia - (Bath et al., 2001). The main objective of the method is to provide a cheap and easy-to-use decision-making tool to small ruminant farmers, in order to allow them to select animals for treatment that are unable to cope with the parasites (usually FamSco 4 or 5), thus leaving untreated animals (usually FamSco 1, 2 or 3) as refugia to delay the spread of anthelmintic-resistant parasites (Van Wyk, 2001).

The designers of the method built the scoring scale and colour chart with reference to the haematocrit or packed cell volume (PCV) of the animals, considered as the "gold standard" (Bath et al., 2001). However, the PCV intervals corresponding to the different FamSco scores are not equal and do not mean the same thing in terms of physiology and pathology. The scoring process lacks precision and the limits between categories are blurred, even when the test is undertaken by trained technicians. Boxplots of PCV values by FamSco show a general trend, but with large overlaps between categories, as reported in Kaplan et al. (2004), Mahieu et al. (2007) and others. Therefore, although the concordance between FamSco scores and anaemia categories derived from PCV measurements is not very good, there are no major consequences for haemonchosis control.

Unfortunately, the choice of numerical category names i.e. FamSco scores ranging from 1 to 5 suggests that it is possible to calculate means and standard deviations, and to make comparisons between means (parametric tests) or medians (non-parametric tests). This is a common pitfall, and many people make this error, especially since the default option of most statistical software packages is to use such data as numerical.

Let consider two fictional groups "a1" and "b1" (100 animals each), with the FamSco distribution reported in contingency Table 1 (X1).

Table 1: distribution of animals according to FamSco ("1" to "5") and to groups ("a1" and "b1")

	"1"	"2"	"3"	"4"	"5"
"a1"	0	10	80	10	0
"b1"	0	32	40	25	3

Using the FamSco as a numerical variable, we can try a t.test, with the Welch correction for unequal variances:

```
welch Two Sample t-test
data: FamSco by group
t = 0.10547, df = 151.95, p-value = 0.9161
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1773155  0.1973155
sample estimates:
mean in group a1 mean in group b1
      3.00          2.99
```

Using a non-parametric test such as the Wilcoxon median comparison test yields a similar result:

```
wilcoxon rank sum test with continuity correction
data: FamSco by group
W = 5165, p-value = 0.6468
alternative hypothesis: true location shift is not equal to 0
median in group a1 median in group b1.
      3              3
```

Conclusion: "a1" and "b1" are drawn from the same population.

Let now consider the same dataset with FamSco as a categorical variable, and perform a Fisher's Exact Test for Count Data. This test is preferred to the basic chi-square test, because some counts (typically from 1 to 5) may produce incorrect chi-square approximations.

```
Fisher's Exact Test for Count Data
data: x1
p-value = 5.293e-08
alternative hypothesis: two.sided
```

The conclusion differs radically, as the Fisher's Exact Test shows that FamSco distribution is not the same for the two groups. From a practical perspective, 28 animals need treatment in the "b1" group instead of only ten in the "a1" group.

Let now consider another example, with the distribution reported in the following contingency Table 2 (X2), and again apply the Welch t-test and the Wilcoxon test.

Table 2: distribution of animals according to FamSco ("1" to "5") and to groups ("a2" and "b2")

	"1"	"2"	"3"	"4"	"5"
"a2"	0	25	70	5	0
"b2"	10	50	25	10	5

welch Two Sample t-test

data: FamSco by group

t = 2.7136, df = 149.41, p-value = 0.007438

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

0.08154847 0.51845153

sample estimates:

mean in group a2 mean in group b2

2.8

2.5

wilcoxon rank sum test with continuity correction

data: FamSco by group

w = 6400, p-value = 0.00019

alternative hypothesis: true location shift is not equal to 0

median in group a2 median in group b2.

3

2

Both Welch t-test and Wilcoxon test suggest that "a2" FamSco values are on average significantly higher than "b2" FamSco values, meaning that "b2" animals are healthier than "a2" ones.

Now, let us run the Fisher's Exact Test for Count Data.

Fisher's Exact Test for Count Data

data: x2

p-value = 1.055e-10

alternative hypothesis: two.sided

Contingency Table 2 shows 15 animals with anaemia in the "b2" group (among them, there are five with severe anaemia) whereas only five require medication in the "a2" group, which is opposite to the previous conclusion, and the Fisher exact test concludes that the differences between these groups are highly significant.

Hence, the conclusions are very clear: considering FamSco as a numerical variable can lead to incorrect conclusions and lead to a failure to detect the most important and useful information — the actual proportion of animals with anaemia or severe anaemia.

Therefore, we must consider the FamSco as an indicator of anaemia, very useful as a decision tool at the farm level, but not as a method for measuring the level of blood haemoglobin or PCV. Using category names such as "very healthy"; "healthy"; "medium"; "anaemic" and "severely anaemic" or for convenience "A"; "B"; "C"; "D" and "E" instead of 1; 2; 3; 4 and 5 could help to avoid the tendency to mistakenly handle the FamSco as numerical data. Moreover, it should also limit the temptation of "precision", where people use .5 or even .25 "score units" (actually implying a ten or twenty category scale). Such precision is unrealistic even in an experimental environment, with the possibility of referring to a gold standard. Under farm conditions or without use of the gold standard, as for other scoring procedures (body condition score etc.) such "precision" is at best a waste of time, and at worst can be a cause of misleading or inappropriate analysis.

References

- Bath, G.F., Hansen, J.W., Krecek, R.C., Van Wyk, J.A., Vatta, A.F. 2001. Sustainable approaches for managing haemonchosis in sheep and goats (FAO), 129 p.
- Bath, G.F., Malan, F.S., Van Wyk, J.A., 1996. The "FAMACHA" ovine anaemia guide to assist with the control of haemonchosis. In: 7th annual congress of the Livestock Health and Production Group of the South African Veterinary Association, Port Elizabeth, June 1996, pp. 152-156.
- Grace, D., Himstedt, H., Sidibe, I., Randolph, T., Clausen, P.H., 2007, Comparing FAMACHA® eye color chart and Hemoglobin Color Scale tests for detecting anemia and improving treatment of bovine trypanosomosis in West Africa. *Vet. Parasitol.* 147, 26-39.
- Kaplan, R.M., Burke, J.M., Terrill, T.H., Miller, J.E., Getz, W.R., Mobini, S., Valencia, E., Williams, M.J., Williamson, L.H., Larsen, M., Vatta, A.F., 2004, Validation of the FAMACHA® eye color chart for detecting clinical anemia in sheep and goats on farms in the southern United States. *Vet. Parasitol.* 123, 105-120.
- Mahieu, M., Arquet, R., Kandassamy, T., Mandonnet, N., Hoste, H., 2007, Evaluation of targeted drenching using Famacha® method in Creole goat: Reduction of anthelmintic use, and effects on kid production and pasture contamination. *Vet. Parasitol.* 146, 135-147.
- Van Wyk, J.A., 2001, Refugia - overlooked as perhaps the most potent factor concerning the development of anthelmintic resistance. *Onderstepoort J. Vet. Res.* 68, 55-67.
- Van Wyk, J.A., Bath, G.F., 2002, The FAMACHA® system for managing haemonchosis in sheep and goats by clinically identifying individual animals for treatment. *Vet. Res.* 33, 509-529.