Tests on Means by Example

Z Test for Comparison of a sample mean with a population mean

Aim: to compare the arithmetic mean of a continuous variable on a representative sample extracted with an known mean (population mean) under the assumption of equality of variance σ^2.

Conditions of application:

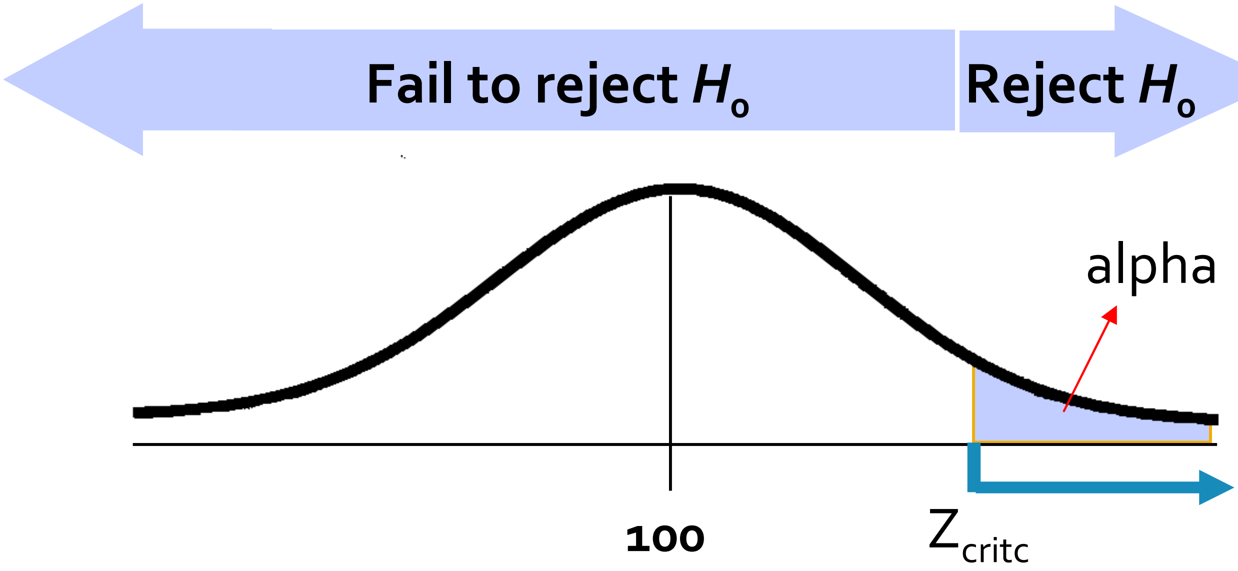
1. The population variation is known and it is equal with the sample variation. (If the variance is not known the Student t test will be applied).
2. The test is correctly applied if the population is normal distributed. Otherwise, the result gives just an orientative response.
3. Sample size higher than or equal to 30.

Right-tailed tests

H_o: μ = 100

H_a/H_1: μ > 100

Points Right



Values that differ "significantly" from 100

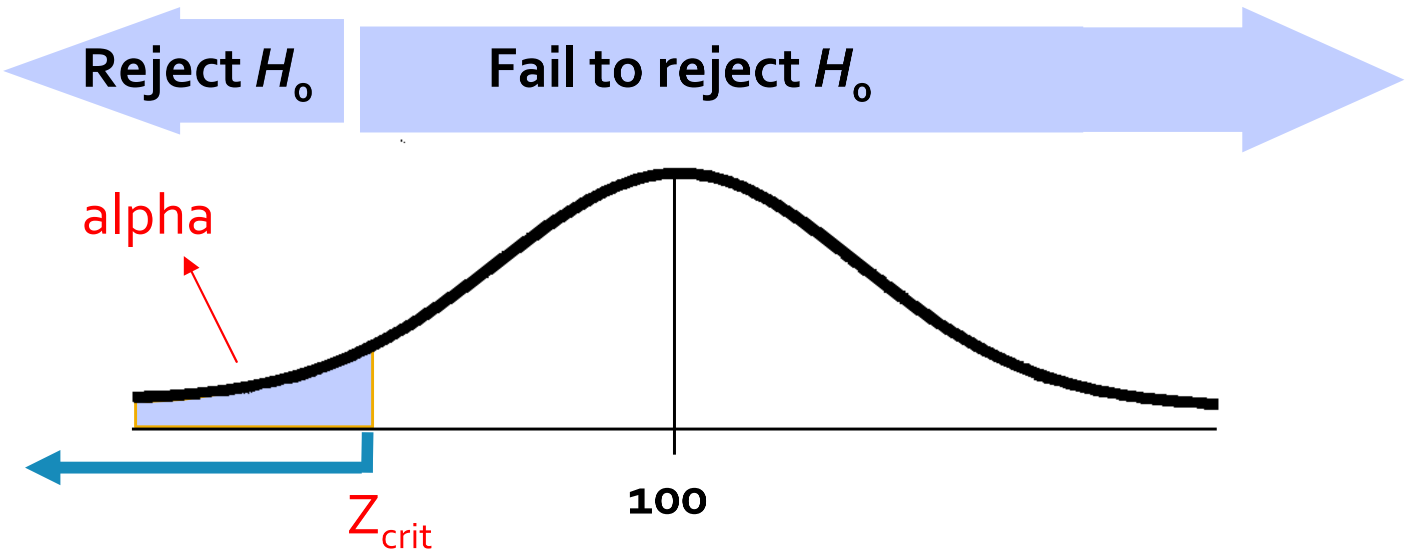
Left-tailed tests

H_o: μ = 100

H_a/H_1: μ < 100

Points Left

Values that differ "significantly" from 100



Z test for comparing a sample mean with a population mean

Hypothesises:

- H_o: the sample and population means are equal
- H_1 (two-tailed test): the sample and population means are NOT equal.

Significance level:

$\alpha = 0.05$.

Critical value (two-tails test)

- $(-\infty , -1.96] \cup [1.96 , \infty)$

- Z parameter:
- $Z = (m - \mu_o)/(\sigma/\sqrt{n})$

Deviation that you got

Deviation from chance alone

- n = sample size
- m = sample mean
- σ = population standard deviation
- μ_o = population mean

T (Student) Test of comparison of a sample means with a known mean (unknown variances)

- **Aim:** to investigate the significance of the difference between a sample and a standardized known mean.
- **Hypothesises**:
  - H_o: the sample and standardized population mean are equal
  - H_a/H_1 (two-tailed test): the sample and standardized population mean are NOT equal.
- **Conditions of application:**
  - The test it is applied when the variance is not known and its estimation on the sample is done on a small sample size which respect the normal distribution assumption. When the assumption of normality is not satisfied than the test loose its validity.
  - If the population variance is known and sample size is larger than 30 the Z test is applied (is a more powerful test).

T (Student) Test of comparison of a sample means with a known mean (unknown variances)

- **Degree of freedom (df):**
  - df = n-1
- **Significance level:**
  - $\alpha = 0.05$
- **Critical interval (two-tailed test):**

$$(-\infty;-t_{n-1,\frac{\alpha}{2}}]\cup[t_{n-1,\frac{\alpha}{2}};+\infty)$$

$$(-\infty;-t_{n-1;0.025}]\cup[t_{n-1;0.025};+\infty)$$

- **T test parameter:**

$$t = (m - \mu_o)/(s/\sqrt{n})$$

  - n = sample size
  - m = sample mean
  - s = sample standard deviation
  - μ_o = population mean

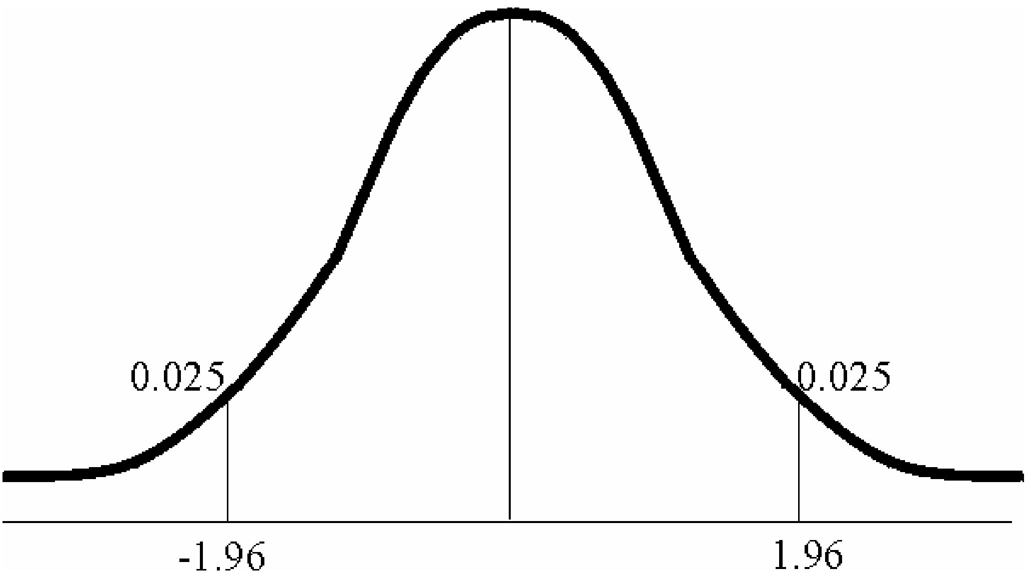
$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n-1}}$$

Z test for comparing two means (known and not equal variances)

- **Aim:** comparing means of quantitative continuous variable when the variances in the two populations are known and un-equal
- **Assumptions***:*
  - The variances of the populations are known. If not, a test of type Student is applied for comparing two means.
  - The data are normal distributed (in both populations). If not, the test give us an orientative value.
- **Null hypothesis**: difference between two means is equal to zero.
- **Alternative hypothesis** for two-tailed test: difference between the two population means is differed by zero.

Z test for comparing two means (known and not equal variances)

- **Significance level** $\alpha = 0.05$.
- **Critical region** for two-tailed test : $(-\infty; -1.96\,] \cup [1.96; \infty)$



- **Statistic**:

$$z = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- m_1 = mean of first sample;
- n_1 = sample size (first sample);
- s_1^2 = variance of first sample;
- m_2 = mean of second sample;
- n_2 = sample size of second sample;
- s_2^2 = variance of second sample.

Student (t) test for comparing two means (unknown and equal variances)

- **Null hypothesis**: Means difference of the two populations is equal to zero.
- **Alternative hypothesis** for two-tailed test: Means difference of the two populations is NOT equal to zero.
- **Assumptions:**
  - The variables in the two samples are normal distributed
  - The variances are equal.
- If these two assumptions are not satisfied the test loss its validity.
- If the variances of populations are known the Z test is applied (is most powerful)

Student (t) test for comparing two means (unknown and equal variances)

- **Degree of freedom (df):**
  - $df = n_1 + n_2 - 2$
- **Significance level**: $\alpha = 0.05$
- **Critical region** for two-tailed test

$$(-\infty; -t_{n_1+n_2-2;\frac{\alpha}{2}}] \cup [t_{n_1+n_2-2;\frac{\alpha}{2}}; +\infty)$$

- **Statistics**

$$t = \frac{m_1 - m_2}{\sqrt{s\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

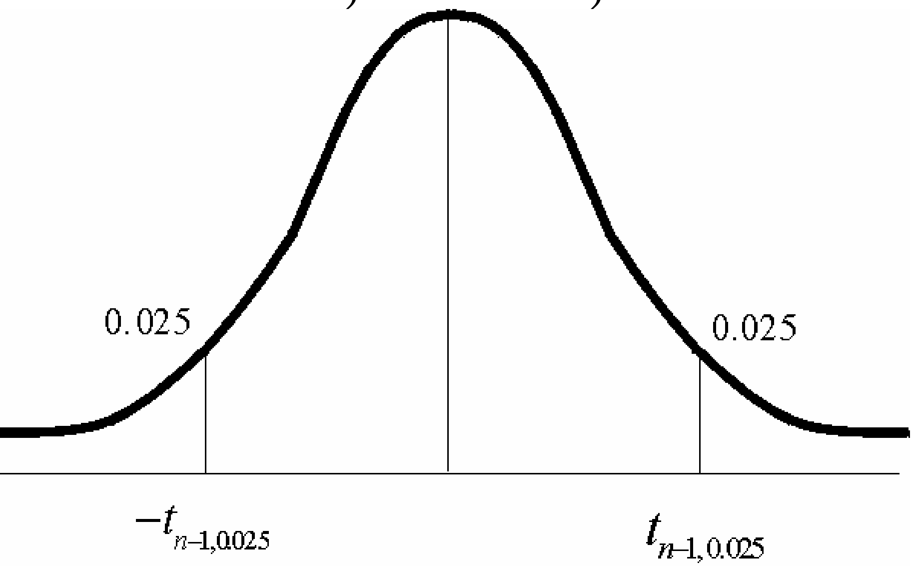
Student (t) for comparing means of paired samples

- **Aim:** comparing the means of two paired samples on quantitative continuous variable (paired means the observation of the same quantitative variable before and after the action of a factor)
- **Assumptions:**
  - Individual observations from the first sample corresponds to a pair in the second sample
  - The differences between pairs of values are normally distributed.
- **Null hypothesis**: The mean of difference of paired data is not significantly different by zero.
- **Alternative hypothesis** for two-tailed test: The mean of difference of paired data is significantly different by zero.

Student (t) for comparing means of paired samples

- **Degrees of freedom (df):** df = n – 1.
- **Significance level**: α = 0.05
- **Critical region:**

$$(-\infty; -t_{n-1;\frac{\alpha}{2}}] \cup [t_{n-1;\frac{\alpha}{2}}; +\infty)$$



- **Statistics**

$$t = \frac{\bar{d}}{\frac{s}{\sqrt{n}}}$$

$$\bar{d} = \frac{(d_1 + d_2 + \ldots + d_n)}{n}$$

- s = standard deviation of differences
- n = sample size

ANOVA test for multiple comparisons

- H_o = the means are equal.
- H_a/H_1 = the means are not equal.

Assumptions:

1. Data are independent by each other.
2. The data of each sample are normally distributed.
3. The standard deviation is the same for all sample.

Remember!

- Attention to the test assumptions!
- If data are quantitative continuous first the normality is verified. Normality tests: Shapiro-Wilk; Kolmogorov-Smirnov; Chi-Square Goodness-of-Fit.
- Comparing two populations means: Z test
- Comparing a sample mean with an known mean: t test
- Comparing more than 2 means: ANOVA
- Comparing two paired means: paired t test.

Attention!!! The statistics for comparing two means is not the same as the statistics for comparing two paired samples.

Tests on Categorical Data I

Testing Association in Contingency Table

- We can perform a hypothesis test on a contingency table. The test we will use most often is the χ2 test.
- χ^2 Test
  - Is proper to be applied if the sample size is large
  - The test is valid if the expected frequency of each cell is at least equal to 1 and the observed frequency is of 5
  - If the above-described conditions are not meet, the Fisher exact test is the proper test

χ^2 Test

- Indicate if that the two variables are or are not independent BUT DO NOT quantify the power of association between them.
- Steps:

1. Define the hypotheses
2. Define the parameter of the test
3. Define the significance level
4. Define the critical interval
5. Calculate the observed value of the parameter of the test
6. Make a decision

χ^2 Test: Problem

- The association between *Streptococcus mutans* (as risk factor) and dental caries was studied. A sample of 620 patients was investigated. The sample contains: 150 patients with caries and *Streptococcus mutans*, 230 patients without caries and without *Streptococcus mutans* and 60 patients with caries but without *Streptococcus mutans.* The presence of *Streptococcus mutans* is asscoiated with dental caries? (df=1; α=0.05; $\chi^2_{critical}$ = 3.84).

χ^2 Test: 1. Hypotheses

- H_o:
  - There is no association between *Streptococcus mutans* and dental caries.
  - The presence of *Streptococcus mutans* and dental caries are independent.
- H_1/H_a:
  - There is an association between *Streptococcus mutans* and dental caries.
  - The presence of *Streptococcus mutans* and dental caries are not independent.

χ^2 Test: 2. Parameter of the test

$$\chi^2 = \sum_{i=1}^{r \cdot c} \frac{(f_i^0 - f_i^t)^2}{f_i^t}$$

Follow a distribution law with (r-1)(c-1) degree of freedom

where

- χ^2 = the parameter of χ^2 test
- f_i^o = observed frequency
- f_i^t = expected (theoretic) frequency

χ^2 Test: 3. Significance level

- Let $\alpha = 0.05$ (5%) be the significance level.

χ^2 Test: 4. Critical region

- Critical region: $[\chi_\alpha^2, \infty)$
- For $\alpha = 0.05$:
  - $\chi_\alpha^2 = 3.84$
  - $[3.48, \infty)$

χ^2 Test: 5. Parameter of the test

observed	DC+	DC-	Total
SP +	TP = 150	FP = 180	330
SP -	FN = 60	TN = 230	290
Total	210	410	620

expected	DC+	DC-	Total
SP +	= 330×210/620	= 330×410/620	330
SP -	= 290×210/620	= 290×410/620	290
Total	210	410	620

χ^2 Test: 5. Parameter of the test

observed	DC+	DC-
SP +	150	180
SP -	60	230

expected	DC+	DC-
SP +	= 112	= 218
SP -	= 98	= 192

$$\chi^2 = \frac{(150-112)^2}{112} + \frac{(180-218)^2}{218} + \frac{(60-98)^2}{98} + \frac{(230-192)^2}{192}$$

$$\chi^2 = \frac{38^2}{112} + \frac{(-38)^2}{218} + \frac{(-38)^2}{98} + \frac{(38)^2}{192}$$

$$\chi^2 = \frac{1444}{112} + \frac{1444}{218} + \frac{1444}{98} + \frac{1444}{192} = 12.89 + 6.63 + 14.73 + 7.52 = \boxed{41.77}$$

χ^2 Test: 6. Making decision

- If $\chi^2 \in [3.84, \infty)$ H_o is rejected with a risk of error of type I (α).
- If $\chi^2 \notin [3.84, \infty)$ H_o is accepted with a risk of error of type II (β).

- Since $41.77 \in [3.84, \infty)$ H_o is rejected with a risk of error of 5%.
- **There is an association between *Streptococcus mutans* and dental caries.**

Continuity correction (Yates's correction)

- For small sample sizes the χ^2 test is too likely to reject the null hypothesis (it tends to spot differences where none really exist).
  - A continuity correction can be made to allow for this.
  - Two conditions have to be met:
    - All expected frequencies must be greater than 1
    - 80% of observed frequencies must be greater than 5

Testul χ²: Corecția Yates

$$\chi^2 = \sum_{i=1}^{r \cdot c} \frac{|f_i^0 - f_i^t|^2 - 0.5}{f_i^t}$$

- 0.5 = Yates correction

Fisher's exact test

- Chi-square procedures can be legitimately applied only if all values of **E** are equal to or greater than 5.
- If a 2×2 contingency table fails to meet the conditions required for the χ^2 test then Fisher's exact test can be used.
- It is based on different mathematics to the χ^2 test which are more robust when sample sizes are small.

Fisher's exact test

- H_o: there is no association between smoking and dental caries
- If the null hypothesis is true - if any ostensible association between smoking and dental caries were the result of nothing more than mere chance coincidence -how likely is it that we might end up with a result this large or larger?

observed	DC+	DC-	Total
smoking +	TP = 2	FP = 7	9
smoking -	FN = 8	TN = 2	10
Total	10	9	19

Fisher's exact test

- Suppose that the initial assessment was performed and the number of subjects who do and do not show characteristics (smoking and dental caries) were counted, but have not yet sorted the subjects according to the correspondences of smoking and dental caries. In this case, all they would have would be the marginal totals shown in the following table/
- Given these marginal totals, there are 10 possible ways in which the specific correspondences between smoking and dental caries.

	DC+	DC-	Total
smoking +			9
smoking -			10
Total	10	9	19

Fisher's exact test

- The p-value is calculated directly from the formula:

$$p = \frac{(a+c)!\,(b+d)!\,(c+d)!\,(a+b)!}{n!\,a!\,b!\,c!\,d!}$$

- The p-value for the observed contingency table must be added to the p-value of the more extreme contingency table.

Fisher's exact test

	DC+	DC-	Total
smoking +	6	2	8
smoking -	1	6	7
Total	7	8	15

	DC+	DC-	Total
smoking +	7	1	8
smoking -	0	7	7
Total	7	8	15

Fisher's exact test

- The p-value must be calculated for the two contingency tables:

$$p_1 = \frac{7!8!7!8!}{15!6!2!6!} = 0.0305$$

$$p_2 = \frac{7!8!7!8!}{15!7!0!7!} = 0.0012$$

- Therefore $p = p_1 + p_2 = 0.0305 + 0.0012 = 0.0317$

Fisher's exact test

- The p-value = 0.0317 < α = 0.05 $\Rightarrow$ that smoking is associated with dental caries.

Summary

- **Conditions for the $\chi 2$ test**
  - All expected values must be greater than 1
  - 80% of expected values must be greater than 5
  - Online calculator
- **r×n: Contingency Table**
  - Online calculator

Tests by Example

- http://www.biomedcentral.com/1746-6148/8/147/abstract
- *BMC Veterinary Research* 2012, **8**:147 doi:10.1186/1746-6148-8-147

- **Background:** Recently, metabolic syndrome (MS) has gained attention in human metabolic medicine given its associations with development of type 2 diabetes mellitus and cardiovascular disease. Canine obesity is associated with the development of insulin resistance, dyslipidaemia, and mild hypertension, but the authors are not aware of any existing studies examining the existence or prevalence of MS in obese dogs.
- Thirty-five obese dogs were assessed before and after weight loss (median percentage loss 29%, range 10-44%). …
- **Results:** Systolic blood pressure ($P = 0.008$), cholesterol ($P = 0.003$), triglyceride ($P = 0.018$), and fasting insulin ($P < 0.001$) all decreased after weight loss, whilst plasma total adiponectin increased ($P = 0.001$). … However, plasma adiponectin concentration was less ($P = 0.031$), and plasma insulin concentration was greater ($P = 0.030$) in ORMD dogs.

Test by Example

- http://www.biomedcentral.com/1746-6148/8/127
- *BMC Veterinary Research* 2012, **8**:127 doi:10.1186/1746-6148-8-127

Background

Enzyme treatment is the mainstay for management of exocrine pancreatic insufficiency (EPI) in dogs. 'Enteric-coated' preparations have been developed to protect the enzyme from degradation in the stomach, but their efficacy has not been critically evaluated. The hypothesis of the current study was that enteric coating would have no effect on the efficacy of pancreatic enzyme treatment for dogs with EPI.

Thirty-eight client-owned dogs with naturally occurring EPI were included in this multicentre, blinded, randomised controlled trial. Dogs received either an enteric-coated enzyme preparation (test treatment) or an identical preparation without the enteric coating (control treatment) over a period of 56 days.

Test by Example

- http://www.biomedcentral.com/1746-6148/8/127
- *BMC Veterinary Research* 2012, **8**:127 doi:10.1186/1746-6148-8-127

Results

There were no significant differences in either signalment or cobalamin status (where cobalamin deficient or not) between the dogs on the test and control treatments. Body weight and body condition score increased in both groups during the trial (*P*<0.001) but the magnitude of increase was greater for the test treatment compared with the control treatment (*P*<0.001). By day 56, mean body weight increase was 17% (95% confidence interval 11-23%) in the test treatment group and 9% (95% confidence interval 4-15%) in the control treatment group. The dose of enzyme required increased over time (*P*<0.001) but there was no significant difference between treatments at any time point (*P*=0.225). Clinical disease severity score decreased over time for both groups (*P*=0.011) and no difference was noted between groups (*P*=0.869). No significant adverse effects were reported, for either treatment, for the duration of the trial.

Test by examples

- http://www.biomedcentral.com/1746-6148/8/153/abstract
- *BMC Veterinary Research* 2012, **8**:153 doi:10.1186/1746-6148-8-153

Background

Dehorning is a common practice involving calves on dairy operations in the United States. However, less than 20% of producers report using analgesics or anesthetics during dehorning. Administration of a systemic analgesic drug at the time of dehorning may be attractive to dairy producers since cornual nerve blocks require 10 -- 15 min to take effect and only provide pain relief for a few hours. The primary objectives of this trial were to (1) describe the compartmental pharmacokinetics of meloxicam in calves after IV administration at 0.5 mg/kg and (2) to determine the effect of meloxicam (n = 6) or placebo (n = 6) treatment on serum cortisol response, plasma substance P (SP) concentrations, heart rate (HR), activity and weight gain in calves after scoop dehorning and thermocautery without local anesthesia.

Results

Plasma meloxicam concentrations were detectable for 50 h post-administration and fit a 2-compartment model with a rapid distribution phase (mean T1/2alpha = 0.22 +/- 0.087 h) and a slower elimination phase (mean T1/2beta = 21.86 +/- 3.03 h). Dehorning caused a significant increase in serum cortisol concentrations and HR (P < 0.05). HR was significantly lower in the meloxicam-treated calves compared with placebo-treated calves at 8 h (P = 0.039) and 10 h (P = 0.044) after dehorning. Mean plasma SP concentrations were lower in meloxicam treated calves (71.36 +/- 20.84 pg/mL) compared with control calves (114.70 +/- 20.84 pg/mL) (P = 0.038). Furthermore, the change in plasma SP from baseline was inversely proportional to corresponding plasma meloxicam concentrations (P = 0.008). The effect of dehorning on laying behavior was less significant in meloxicam-treated calves (p = 0.40) compared to the placebo-treated calves (P < 0.01). Calves receiving meloxicam prior to dehorning gained on average 1.05 +/- 0.13 kg bodyweight/day over 10 days post-dehorning compared with 0.40 +/- 0.25 kg bodyweight/day in the placebo-treated calves (p = 0.042).

Tests by Example

- http://www.biomedcentral.com/1746-6148/7/58
- *BMC Veterinary Research* 2011, **7**:58 doi:10.1186/1746-6148-7-58

Table 3

Clinicopathologic characteristics of canine mammary carcinoma

	KLF4 expression (Quick score)					
	low/moderate (< 9)		**High (≧ 9)**		**N**	**P**
Survival[a]						
< 24 months	9	56.3%	34	89.5%	43	0.010
≧ 24 months	7	43.8%	4	10.5%	11	

[a] Twenty-one cases lacked survival data and were excluded from the analysis.

Chu *et al. BMC Veterinary Research* 2011 **7**:58 doi:10.1186/1746-6148-7-58

Test by examples

- http://www.biomedcentral.com/1746-6148/8/18
- *BMC Veterinary Research* 2012, **8**:18 doi:10.1186/1746-6148-8-18

Table 1

Composition of white blood cells during acute ASFV infection in swine.

Cell types	The percent (%) of cells								ANOVA
	Control	1 dpi	2 dpi	3 dpi	4 dpi	5 dpi	6 dpi	7 dpi	p
Lymphoblasts	0	0	0.5 ± 0.1	3.2 ± 0.9**	3.6 ± 0.9**	18.9 ± 4.8**	10.3 ± 2.7**	2.8 ± 0.7**	< 0.001

Control represents the mean of given cell type for each time point of infection

* Significant decrease compared with control ($p < 0.05$-$p < 0.001$)

** Significant increase compared with control ($p < 0.05$-$p < 0.001$)

Karalyan *et al*. *BMC Veterinary Research* 2012 **8**:18   doi:10.1186/1746-6148-8-18