



Se EU-Kommissionen, Den Europæiske Landbrugsfond for Udvikling af Landdistrikterne

## Note on the systematic use of differences in functional traits when simulating differences in functional traits, for different management levels (=herds).

Jehan Ettema, MBG, AU,

The reason this note is written, is the uncertainty concerning the representation of differences in functional traits when simulating cross-breeding. The question is whether using absolute differences in e.g. disease occurrence reported in scientific literature (calf mortality in pure bred of 15% versus 7% in crossbreds, a difference of 8%-point) should also be used, in case the potential for crossbreeding should be studied in a herd where calf mortality with pure bred animals today is only 5%. Alternatively, odds ratios (OR) or risk ratios (RR), measures often used quantities in epidemiology, should be used to represent the difference in disease risk, mortality risk and conception rates in herd with different management levels.

### Contents

Summary .....	1
Calculating the OR and RR given difference in incidences for a Treatment and Control group .....	2
Using OR, RR or absolute differences to calculate probability for Treatment .....	4
Data from SEGES on differences in reproductive performance .....	5
Number of inseminations per pregnancy .....	5
Interval first and last insemination .....	6
Mastitis treatments .....	8

### Summary

When using absolute difference to express the difference between crossbred cows and purebred cows, it is possible that the calculated probability for the crossbred cows ends up below zero. This is from a principal point of view undesirable. When using a risk ratio, it is possible that the calculated probability for the crossbred cows exceeds 1. This is from a principal point of view undesirable. However, for disease risks calculated for weekly steps (very low values), the weekly risk would not exceed 1, but the lactational incidence could exceed one. More than 1 case during a lactation is not an exception. When using an odds ratio to express the difference between crossbred cows and

purebred cows, the calculated probability will always be between 0 and 1. The effect of an odds ratio is largest around the probability of 0,5 and decreases towards 0 and 1.

Data from seges on the difference between cows with different breeding values for reproductive traits in herds with different management levels do not support any decision on using absolute values, risk ratios or odds ratios. It was found however, that even though differences in an interval between first and last insemination are the same, the pregnancy rates *behind* these differences are not the same.

Data from seges on the difference between cows with different breeding values for mastitis in herds with different management levels does support using risk ratios or odds ratios. Based on this data, use absolute differences seems incorrect.

### Calculating the OR and RR given difference in incidences for a Treatment and Control group

As a start, the difference between OR and RR is presented.

	Diseased	Healthy	
Treatment (exposed)	$D_E$	$H_E$	$N_E$
Control (not-exposed)	$D_{NE}$	$H_{NE}$	$N_{NE}$

$$RR = \frac{D_E/N_E}{D_{NE}/N_{NE}},$$

$$OR = \frac{D_E/H_E}{D_{NE}/H_{NE}},$$

An illustration of differences between the resulting RR's and OR's for three values for Treatment and 20 levels of Control is presented in table 1.

Table 1: Odds Ratio (OR) and Risk Ratio (RR) calculated for the Treatment-group compared to the Control group, where the incidence of a certain disease in the treatment group is 0,5, 0,9 and 0,1 and the incidence of the disease in the control group varies between 0 and 1.

Treatment= 0,5			Treatment= 0,9			Treatment= 0,1		
Control	OR	RR	Control	OR	RR	Control	OR	RR
0,05	19,00	10,00	0,05	171,00	18,00	0,05	2,11	2,00
0,1	9,00	5,00	0,1	81,00	9,00	0,1	1,00	1,00
0,15	5,67	3,33	0,15	51,00	6,00	0,15	0,63	0,67
0,2	4,00	2,50	0,2	36,00	4,50	0,2	0,44	0,50
0,25	3,00	2,00	0,25	27,00	3,60	0,25	0,33	0,40
0,3	2,33	1,67	0,3	21,00	3,00	0,3	0,26	0,33
0,35	1,86	1,43	0,35	16,71	2,57	0,35	0,21	0,29
0,4	1,50	1,25	0,4	13,50	2,25	0,4	0,17	0,25
0,45	1,22	1,11	0,45	11,00	2,00	0,45	0,14	0,22
0,5	1,00	1,00	0,5	9,00	1,80	0,5	0,11	0,20
0,55	0,82	0,91	0,55	7,36	1,64	0,55	0,09	0,18
0,6	0,67	0,83	0,6	6,00	1,50	0,6	0,07	0,17
0,65	0,54	0,77	0,65	4,85	1,38	0,65	0,06	0,15
0,7	0,43	0,71	0,7	3,86	1,29	0,7	0,05	0,14
0,75	0,33	0,67	0,75	3,00	1,20	0,75	0,04	0,13
0,8	0,25	0,63	0,8	2,25	1,13	0,8	0,03	0,13
0,85	0,18	0,59	0,85	1,59	1,06	0,85	0,02	0,12
0,9	0,11	0,56	0,9	1,00	1,00	0,9	0,01	0,11
0,95	0,05	0,53	0,95	0,47	0,95	0,95	0,01	0,11
1	0,00	0,50	1	0,00	0,90	1	0,00	0,10

The lower the incidences, the more does the OR equal the RR, as can also be concluded from the equations on page 1. Figure 1 presents the results from table 1, for the Treatment group with the incidence of mastitis at 0,5. Both the OR and RR are 1, when the incidence for Control equals the incidence for Treatment of 0,5. The OR has higher values compared to the RR above 1, when  $C < T$ . Values for OR are closer to zero compared to the RR below 1, when  $C > T$ .

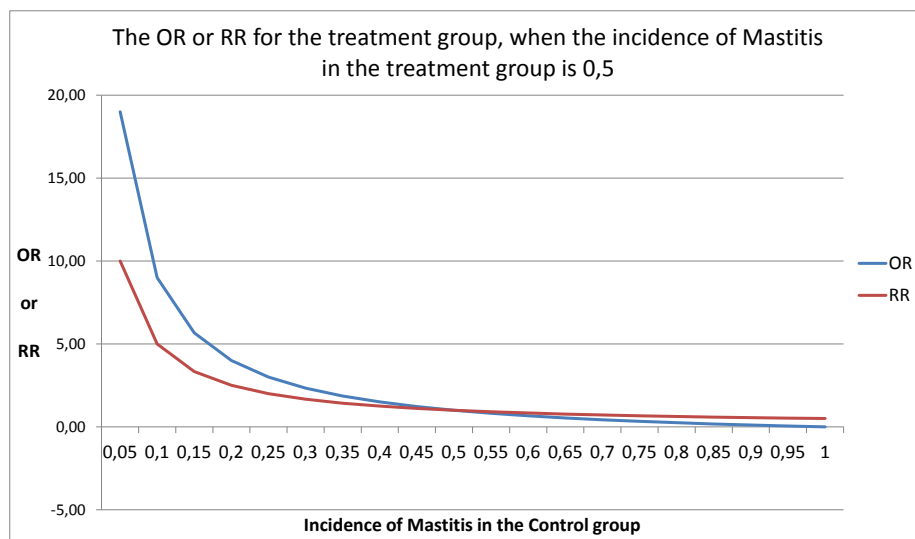


Figure 1: OR and RR calculated for the Treatment-group compared to the Control group, where the incidence in the treatment group is 0,5 and the incidence in the control group varies between 0 and 1.

## Using OR, RR or absolute differences to calculate probability for Treatment

In case we have an incidence of 40% for Control (C) and 20% for Treatment (T), the difference can be expressed as an absolute difference of -20% (-0,2), an OR of 0,375 or a RR of 0,5. The consequences of using the absolute difference, OR and RR to calculate the risk for T given other levels of disease incidences for C (below and above 40%) is presented in figure 2.

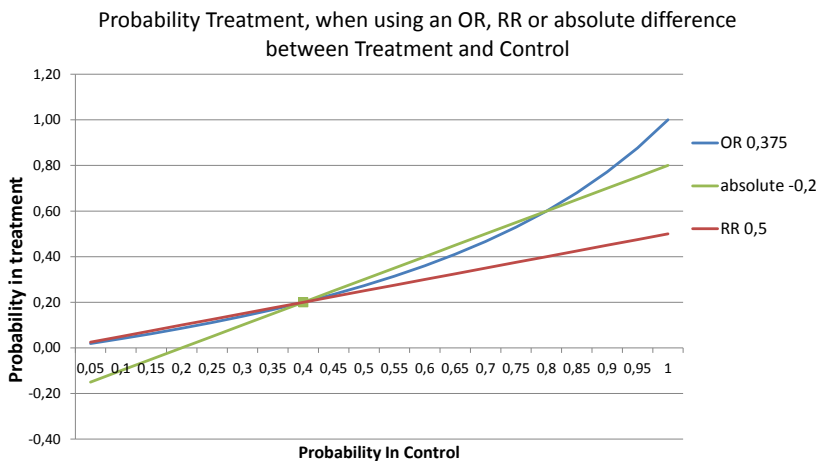


Figure 2: calculated probability for Treatment for different levels of probability for Control and when using an RR of 0,5, an OR of 0,375 and an absolute difference of -0,2.

In figure 3 the relationship between risk for Treatment and Control is presented, when the difference between Control (40%) and Treatment (60%) is positive and the OR and RR are 2,25 and 1,5 respectively.

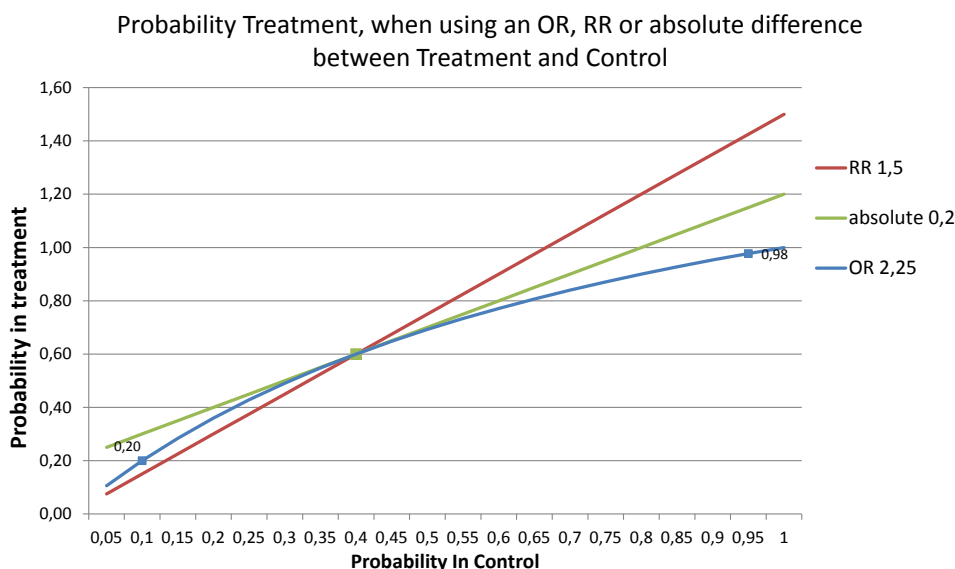


Figure 3: calculated probability for Treatment for different levels of probability for Control and when using an RR of 1,5, an OR of 2,25 and an absolute difference of +0,2.

In case the difference between T and C is negative (and the OR and RR < 1), using absolute differences to calculate the probability for T results in negative probabilities as soon as the probability for C is smaller than the difference (for probabilities for C below 0,2 in figure 2). When using RR, the probability for T can't exceed 0,5. When using OR the probability for T can be anywhere between 0 and 1.

In case the difference between T and C is positive (and the OR and RR >1), using absolute differences to calculate the probability for T can result in probabilities larger than 1 (for probabilities for C above 0,8 in figure 3). When using RR, the probability for T can also exceed 1. When using OR the probability for T can be anywhere between 0 and 1. A probability of 0,1 for C results in a probability of 0,2 for T. However, a probability of 0,95 for C, only results in a slightly higher probability of 0,98 for T, while the difference between them still equals an OR of 2,25. **The effect of an OR decreases towards a probability of 1.**

## Data from SEGES on differences in reproductive performance

A population of herds has been divided into good herds (god besætning) and bad herds (dårlig besætning) for a number of reproduction traits. Within each herd, the value for the traits is presented for the cows with a breeding value lower (lav avl) than the population average and higher than the population average (høj avl). Results are presented in table 3. The *middel herd* has been deleted from the data presented in table 3. So the bad herd is the 33% worst herds and the good is the 33% best performing herds.

### Number of inseminations per pregnancy

Table 2 presents data from SEGES' data (Ane Closter, forskelsanalyse af management niveauer hos besætninger med renracet Holstein køer, table 7). The data for female fertility is expressed as *number of insemination per pregnancy* (antal ins) for different parities. Furthermore, an approximation of conception rate is calculated based on the number of inseminations (1/antal ins). This approximation is relevant, since SimHerd needs to work with a conception rate rather than number of insemination. Differences (forskell) are calculated for "antal ins" and for "1/antal ins". OR and RR are calculated for "1/antal ins".

Table 2: Antal insemineringer for kvier, 1. og 2 kalvskøer opdelt i lav, middel [ikke vist i tabellen] og Høj fænotypisk niveau indenfor besætning. Data er opdelt Lav (frugtbarhed), Middel (frugtbarhed) [ikke vist i tabellen] og Høj (frugtbarhed) genetisk niveau indenfor besætning. Den fænotypiske opdeling af besætning managementt er baseret på summen af "Interval mellem 1. og sidste inseminering for 1. og 2 kalvskøer.

	Fænotyp. Lav		Fænotyp. Høj		Note
	Gen. Lav	Gen. Høj	Gen. Lav	Gen. Høj	
<b>antal ins, kvie</b>	1,74	1,70	1,61	1,58	
- forskel	-0,04		-0,03		
<b>1/antal ins ≈ conception rate</b>	0,57	0,59	<b>0,62</b>	<b>0,63</b>	
- forskel	0,014		0,012		
- RR	1,02		1,02		
- OR	1,06		1,05		
<b>antal ins, 1. kalvs</b>	2,36	2,21	1,98	1,85	
- forskel	-0,15		-0,13		a
<b>1/antal ins ≈ conception rate</b>	0,42	0,45	<b>0,51</b>	<b>0,54</b>	
- forskel	0,029		0,035		b
- RR	1,07		1,07		c
- OR	1,12		1,15		d
<b>antal ins, 2. kalvs</b>	2,45	2,37	2,06	2,91	e
- forskel	-0,08		0,85		
<b>1/antal ins ≈ conception rate</b>	0,41	0,42	<b>0,49</b>	<b>0,34</b>	
- forskel	0,014		-0,142		
- RR	1,03		0,71		
- OR	1,06		0,55		

Noter:

- The difference in "antal ins" is largest in the poor performing herds (Fænotypisk lav).
- The difference in *conception rate* is largest for the good performing herds (Fænotypisk høj)
- RR is the same for the two herds
- OR is largest for the good herds.
- Are these results correct??? Gen. Høj has a much higher number of ins. Ask Ane to verify.

Based on these results from SEGES and the derived estimates for conception rate, it can't be concluded that justifies the use of absolute differences, OR or RR for this trait.

## Interval first and last insemination

Table 3 shows the results from the same data set on the interval between first and last insemination (Table 5 from Ane Closter). It is difficult to conclude whether the differences between cows with a low and high breeding value are constant, when comparing these cows with two breeding values at the two management levels. The differences are significant (see Ane's report), but the magnitude of the differences is small.

Table 3: Interval mellem 1. og sidste inseminering for kvier, 1. og 2 kalvskøer opdelt i lav, middel\* og Høj fænotypisk niveau indenfor besætning. Data er opdelt Lav (frugtbarhed), Middel (frugtbarhed)\* og Høj (frugtbarhed) genetisk niveau indenfor besætning. Den fænotypiske opdeling af besætning managementt er baseret på summen af "Interval mellem 1. og sidste inseminering for 1. og 2 kalvskøer.

\* værdien for middel er ikke vist i nedenstående tabel, men værdien var med i den oprindelige tabel fra Ane.

	Fænotyp. Lav		Fænotyp. Høj	
	Gen. Lav	Gen. Høj	Gen. Lav	Gen. Høj
<b>Interval 1. og sidste ins, kvie</b>	27,2	24,9	20,6	18,9
- forskel	-2,3		-1,7	
<b>Interval 1. og sidste ins, 1. kalvs</b>	66,5	55,0	44,1	35,5
- forskel	-11,5		-8,7	
<b>Interval 1. og sidste ins, 2. kalvs</b>	73,7	64,1	49,8	39,5
- forskel	-9,6		-10,3	

**A note should be made to these differences:** The interval between first and last insemination is the result of heat detection rate and conception rate. The product of these two rates is the 21-d-pregnancy rate or reproductive efficiency (read more on page 8).

The same interval, does not mean, that the difference in 21-d-pregnancy rate is the same

- An interval reduction from **44 days to 34 days (-10)** is the result of an increase in pregnancy rate from 0,20 to 0,263 (**+0,063**). This corresponds to a RR of 1,32 and an OR of 1,43.
- An interval reduction from **65 days to 55 days, (-10)** is the result of an increase in pregnancy rate from 0,15 to 0,185 (**+0,035**). This corresponds to a RR of 1,23 and an OR of 1,29.

The results show that even though the difference in interval between two levels of breeding values for reproductive performance are the same, the absolute differences in 21-d-pregnancy rate are large for herds with good and bad reproduction.

This argues against using absolute differences for pregnancy rate. Whether the OR's are the same (1,43 and 1,29) and therefore the correct quantity to use, can be argued.

The relationship between pregnancy rate and interval between first and last insemination is based on figure 4. In figure 4 the number of cows getting pregnant is plotted against the number of days after first insemination.

With a pregnancy rate of 0,20, the conception rate is assumed to be 45% and the heat observation is also 45%. When breeding 1000 cows, 450 will get pregnant and the number of days between first and last insemination is 0 for these 440 cows. Twenty-one-days later, the remaining 550 cows can be bred, but only 45% (247) gets observed in heat and 45% of them (111) get pregnant. The number of days between first and last insemination is 21 for these 111 cows. In the same way the number of cows can be calculated that have an interval of 42, 63, ..., 231 days. Accordingly, the weighted average for the intervals is calculated  $(450 \times 0 + 111 \times 21 + \dots) / (450 + 111 + \dots) = 44$ . After 231 days, 95% of the initial 1000 cows are pregnant.

With a pregnancy rate of 0,15, the same weighted average can be calculated. For this lower rate however, 95% of the initial 1000 cows are pregnant after 315 days instead of 231 days. The

assumption on how many days should be included in this calculation also affects the calculated values for the interval between first and last insemination.

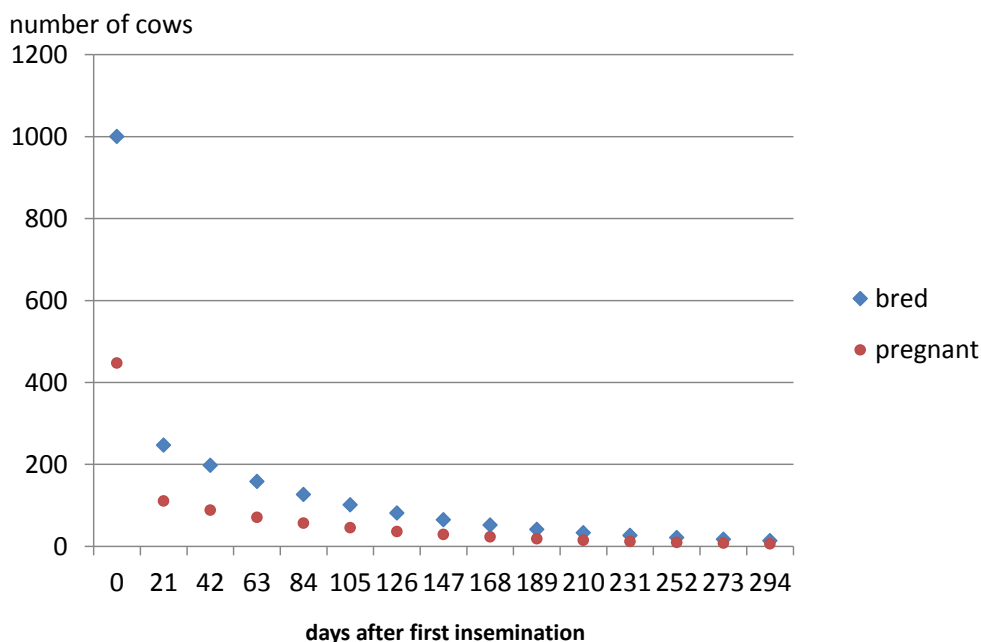


Figure 4: number of cows getting bred and pregnant after insemination when 21-d-pregnancy rate is 0,2.

## Mastitis treatments

Table 4 shows the results from the same data set on the interval between first and last insemination (Table 9 from Ane Closter).

Table 4: Den genetiske niveau opdeling indenfor besætning er baseret på Mange, Middel\* og Få mastitis tilfælde samt Mange, Middel\* og Få mastitis tilfælde. Den fænotypiske opdeling af besætning managementt er baseret på summen af mastitis tilfælde per besætning for 1. og 2. kalvskøer og data er op delt i Mange, Middel og Få mastitis tilfælde.

\* værdien for middel er ikke vist i nedenstående tabel, men værdien var med i den oprindelige tabel fra Ane.

	Fænotyp. Lav		Fænotyp. Høj	
	Gen. Lav	Gen. Høj	Gen. Lav	Gen. Høj
<b>mast i 1. lakt</b>	0,32	0,19	<b>0,12</b>	<b>0,07</b>
- forskel	-0,131		-0,051	
- RR	0,59		0,56	
- OR	0,50		0,53	
<b>mast i 2. lakt</b>	0,39	0,32	<b>0,10</b>	<b>0,08</b>
- forskel	-0,071		-0,019	
- RR	0,82		0,81	
- OR	0,73		0,80	



The results are clear for mastitis treatments. The absolute differences between genetically good and bad cows differ, depending on the management level. As soon as differences are expressed as either OR or RR, the differences are the same. Table 6 shows the OR's for other contrasts.

Table 6: odds ratios calculated for the different combinations: middel and få compared to mange. And få compared to middel. Data based on table 9 from Ane.

1. lakt	Fænotyp. Opdeling	Gen. Opdeling	behandlinger	odds	OR, mange er ref.	OR, middel er ref.
		mange	0,32	0,471		
	mange	middel	0,251	0,335	0,71	
		få	0,189	0,233	0,50	0,70
		mange	0,205	0,258		
	middel	middel	0,155	0,183	0,71	
		få	0,116	0,131	0,51	0,72
		mange	0,116	0,131		
	få	middel	0,081	0,088	0,67	
		få	0,065	0,070	0,53	0,79
2. lakt						
		mange	0,388	0,634		
	mange	middel	0,317	0,464	0,73	
		få	0,27	0,370	0,58	0,80
		mange	0,232	0,302		
	middel	middel	0,206	0,259	0,86	
		få	0,171	0,206	0,68	0,80
		mange	0,132	0,152		
	få	middel	0,102	0,114	0,75	
		få	0,083	0,091	0,60	0,80