

## 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, 02-05-2016

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 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), quantities typically used 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 .....                | 3 |
| Data from SEGES on differences in reproductive performance .....                                 | 4 |
| Mastitis treatments .....                                                                        | 4 |

### 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 one 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 udder health in herds with different management levels supports the decision to use odds ratios when calculating different levels of disease risk. Based on this data, using absolute differences seems incorrect.

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.

## 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 below.

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

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

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

### Example:

|           | Diseased | Healthy |     |
|-----------|----------|---------|-----|
| Treatment | 2        | 98      | 100 |
| Control   | 4        | 96      | 100 |

$$RR = (2/100)/(4/100) = 0,5$$

$$OR = (2/98)/(4/96) = 0,49$$

In case of a low incidence of a disease, the values for N (100) and H (96 and 98) are almost the same, which results in the OR and the RR to be almost the same.

### Example:

|           | Diseased | Healthy |     |
|-----------|----------|---------|-----|
| Treatment | 30       | 70      | 100 |
| Control   | 60       | 40      | 100 |

$$RR = (30/100)/(60/100) = 0,5$$

$$OR = (30/70)/(60/40) = 0,29$$

In case of a high incidence of a disease, the values for N (100) and H (70 and 40) are further apart, which results in the OR and the RR being further apart.

## 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)**, or with an **OR of 0,375** or with a **RR of 0,5**.

The consequences of using the absolute difference, OR and RR to calculate the probability for T given other levels of disease incidences for C (below and above 40%) is presented in figure 1.

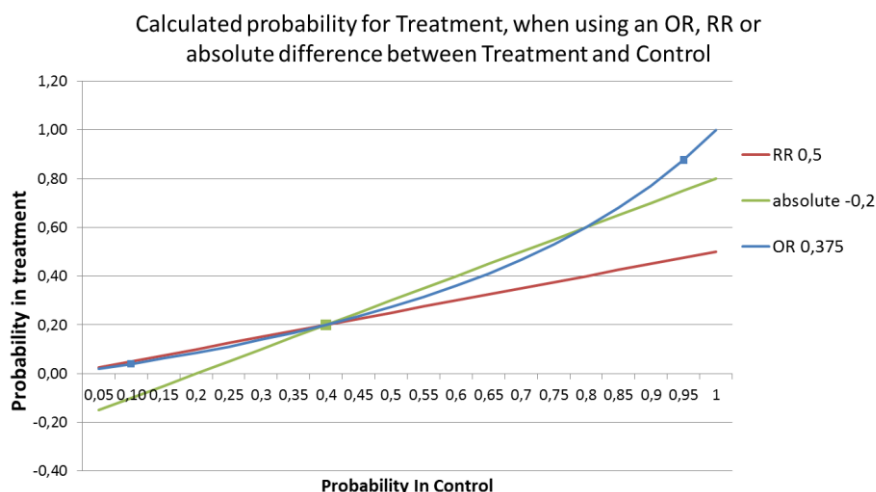


Figure 1: calculated probability for Treatment for different levels of probability for Control and when using an RR of 0,5, an OR of 0,375 and an absolute difference of -0,2.

In figure 2 the relationship between probability for Treatment and Control is presented, when the absolute **difference between Control (40%) and Treatment (60%) is +20%** and the **OR and RR are 2,25 and 1,5** respectively.

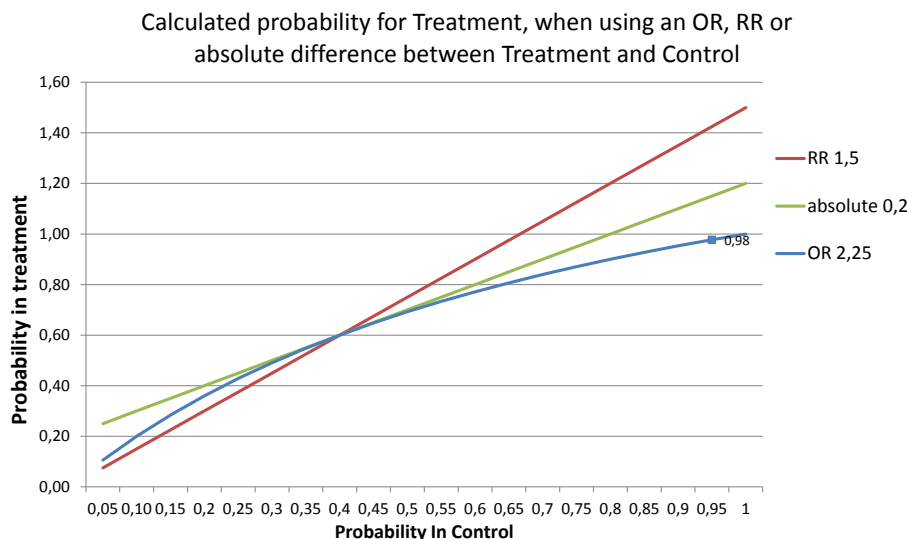


Figure 2: calculated probability for Treatment for different levels of probability for Control and when using an RR of 1,5, an OR of 2,25 and an absolute difference of +0,2.

In case the difference between T and C is negative (and the OR and RR < 1), using the absolute difference to calculate the probability for T results in negative probabilities for low values of C (figure 1). When using the RR, the probability for T can't exceed 0,5 for values of C between 0 and 1. When using OR the calculated 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 high values of C. 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 0 and 1; *the effect is largest in the middle.*

## Data from SEGES on differences in fertility and disease occurrence

Closter and Fogh: Analyse af managementniveau hos renracet Holstein besætninger.

### Mastitis treatments

Table 1: A population of herds has been divided into herds with good management (Fænotyp. Høj) and bad management (Fænotyp. Lav) for the sum of mastitis cases (Høj management = low number of cases). Within the herds, cows are divided in groups with the lowest (Gen. Lav) and highest genetic level (Gen. Høj) for mastitis cases. Values for the herds with average management (middle) and cows with average genetic level are omitted, but those values were present in [table 9](#) of Closter and Fogh. With the middle-group missing, the values for high and low are representing the best and worst 33% of the herds and the cows.

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

\* Cases of mastitis, values from Closter and Fogh

\*\* calculated values

The absolute differences between genetically good and bad cows differ, depending on the management level (-0,131 vs. -0,051 (for mastitis in first lactation)). When expressing differences as either OR (0,50 vs. 0,53) or RR (0,59 vs. 0,56), the differences are the same. Table 2 shows the OR's for contrasts where the middle group is also included. The same conclusion can be drawn from this table.

Table 2: A population of herds has been divided into herds with good management (få), average (middel) and bad management (mange) for the sum of mastitis cases (få tilfælde = good management). Within the herds, cows are divided (Gen. opdeling) in groups of low, average and high genetic level (mange, middel and få, respectively). Odds ratios are calculated within the different phenotypic levels for the different contrast for "Gen. opdeling": middel and få compared to mange and få compared to middel. Data based on table 9 from Closter and Fogh.

| 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                 |

\* values from Closter and Fogh

\*\* calculated values

## Number of inseminations per pregnancy

Table 3 presents data from SEGES' data (Closter and Fogh, 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 (forskøl) are calculated for "antal ins" and for "1/antal ins". OR and RR are calculated for "1/antal ins".

Table 3: A population of herds has been divided into herds with good management (Fænotyp. Høj) and bad management (Fænotyp. Lav) for the number of inseminations (Høj management = low number of inseminations). Within the herds, cows are divided in groups with the lowest (Gen. Lav) and highest genetic level (Gen. Høj) for number of inseminations. Values for the herds with average management (middle) and cows with average genetic level are omitted, but those values were present in [table 7](#) of Closter and Fogh. With the middle-group missing, the values for high and low are representing the best and worst 33% of the herds and the cows.

|                                        | 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     | a    |
| - forskøl*                             | -0,04        |          | -0,03        |          |      |
| <b>1/antal ins ≈ conception rate**</b> | 0,57         | 0,59     | 0,62         | 0,63     |      |
| - forskøl**                            | 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     | a    |
| - forskøl*                             | -0,15        |          | -0,13        |          |      |
| <b>1/antal ins ≈ conception rate**</b> | 0,42         | 0,45     | 0,51         | 0,54     |      |
| - forskøl**                            | 0,029        |          | 0,035        |          |      |
| - RR**                                 | 1,07         |          | 1,07         |          |      |
| - OR**                                 | 1,12         |          | 1,15         |          |      |
|                                        |              |          |              |          |      |
| <b>antal ins, 2. kalvs*</b>            | 2,45         | 2,37     | 2,06         | 1,91     | b    |
| - forskøl*                             | -0,08        |          | -0,15        |          |      |
| <b>1/antal ins ≈ conception rate**</b> | 0,41         | 0,42     | 0,49         | 0,52     |      |
| - forskøl**                            | 0,014        |          | 0,038        |          |      |
| - RR**                                 | 1,03         |          | 1,08         |          |      |
| - OR**                                 | 1,06         |          | 1,16         |          |      |

\* values for number of inseminations, from Closter and Fogh

\*\* calculated values

Note:

- For "kvie" and "1. kalvs" the differences expressed in absolute differences, OR and RR are more or less the same when comparing cows with high and low genetic levels in different herds.
- For "2. kalvs", the differences between cows with the lowest and highest genetic levels are largest in well performing herds. The differences are largest regardless of how they are expressed (absolute for number on insemination, absolute number for conception rate and OR or RR for conception rate).

Based on these results from SEGES and the derived estimates for conception rate, it can't be concluded whether using absolute differences, OR or RR is most appropriate for this trait.