

DIRK HINRICHS, ECKHARD STAMER, WOLFGANG JUNGE and ERNST KALM

Genetic Analysis of Several Disease Categories Using Test Day Threshold Models in German Holstein Cows

Abstract

In the present study trajectories for the disease categories udder diseases, metabolic diseases, all diseases, fertility diseases, and ovarian problems, are described. Variance components were estimated where the considered period of lactation varies between 50 and 300 days. Furthermore, the impact of the number of daughters per sire was analysed with disease information from the first 50 days of lactation. In total 18 data sets were analysed with a test day threshold model. The average disease frequencies were between 6.5% and 2.7% for udder diseases and they were between 1.7% and 0.4%, and 15.3% and 6.0% for metabolic diseases and 'all diseases', respectively. For udder diseases the estimated heritabilities vary between 0.12 and 0.25 depending on lactation length and the number of daughters per sire. For metabolic disease heritabilities were estimated within the interval of 0.12 to 0.24. Depending on the lactation length the heritabilities of all diseases were fairly constant (0.03 to 0.04). The heritabilities of all diseases increased if fertility diseases were excluded. Those estimates vary between 0.15 to 0.19.

Key Words: Threshold models, disease categories, variance components

Zusammenfassung

Titel der Arbeit: Genetische Analyse verschiedener Erkrankungskategorien mit Testtagsschwellenwertmodellen bei deutschen Holsteins

In der vorliegenden Untersuchung werden die Laktationsverläufe der Erkrankungskategorien Eutererkrankungen, Stoffwechselerkrankungen, Fruchtbarkeits- und Zyklusstörungen sowie alle Erkrankungen beschrieben. Anschließend wurden Varianzkomponenten mit einem Schwellenwert Testtagsmodell geschätzt. Die Laktationslänge variierte zwischen 50 und 300 Tagen. Für die ersten 50 Laktationstage wurde der Einfluss der Anzahl der Töchter (verbesserte genetische Struktur) untersucht. Insgesamt wurden 18 verschiedene Datensätze analysiert. Die mittlere Frequenz der Eutererkrankungen lag zwischen 2,7 und 6,5 %. Bei den Stoffwechselerkrankungen variierte dieser Wert zwischen 0,4 und 1,7 %. Für die Kategorie alle Erkrankungen zusammen betrachtet lagen die Frequenzen zwischen 6,0 und 15,3 %. Die Heritabilitäten der Eutererkrankungen schwankten zwischen 0,12 und 0,25, und die der Stoffwechselerkrankungen variierten zwischen 0,12 und 0,24. Für alle Erkrankungen lagen die Heritabilitäten zwischen 0,03 und 0,04. Höhere Heritabilitäten wurden geschätzt, nachdem die Fruchtbarkeitserkrankungen von der Analyse ausgeschlossen wurden. Diese Schätzwerte variierten zwischen 0,15 und 0,19.

Schlüsselwörter: Schwellenwertmodelle, Erkrankungskategorien, Varianzkomponenten

Introduction

Milk production has changed during the past two decades and has been characterised by increasing herd size, and increasing milk yield, as well as increasing veterinary costs (e.g. JAKOB and DISTL, 1998). Today, diseases (e.g. fertility and udder health problems) are one of the most significant problems affecting commercial milk production. Diseases reduce animal welfare and result in economic losses for the farmer due to extra veterinary treatments, extra labour, decreasing milk production, discarded milk and involuntary early culling (NIELSEN et al., 1999). The frequencies of

diseases or disease categories depend on case definition and show considerable variation in literature reports (KELTON et al., 1998; COLLARD et al., 2000). Most analyses of disease data were carried out using linear models and disease information was considered as ‘all or none’ trait. These models apply a normal distribution of the data. Threshold models are an alternative to the traditional linear models as they take the binary nature of disease data into account (GIANOLA and FOULLEY, 1983). The estimates for heritabilities are higher when lactation threshold models are used (SIMIANER et al., 1991; HERINGSTAD et al., 1997; HINRICHS et al., 2003; HINRICHS et al., 2004). However, it should be noted that the definition of disease information as ‘all or none’ trait does not utilise all disease information provided by the data, since some cows have more than one disease case per lactation. REKAYA et al. (1998) therefore recommended the development of test day models for the analysis of longitudinal binary response such as disease field data. The use of longitudinal threshold models would improve the estimation of genetic parameters, because repeated disease cases are distinct observations in those models.

Heritabilities for some metabolic diseases has been estimated by SIMIANER et al. (1991), LYONS et al. (1991), URIBE et al. (1995), and NIELSEN et al. (1999) and the estimated heritabilities vary between 0.01 and 0.47, with most estimates in the interval of 0.08 to 0.15. Most literature reports of heritabilities for other diseases are based on dividing all available disease information into two disease categories, mastitis and other diseases than mastitis, respectively. Estimates for diseases other than mastitis (0.11) are given by LUND et al. (1994) and HANSEN et al. (2002), who provided a heritability of 0.02.

If heritabilities for udder diseases (e.g. mastitis) are estimated with linear models most values are in the interval of 0.02 to 0.03 (HERINGSTAD et al., 2000). Some literature reports could be found where mastitis data were analysed with test day threshold models (REKAYA et al., 1998; KADARMIDEEN et al., 2000; HERINGSTAD et al., 2003b). The resulting estimates for the heritabilities ranged between 0.03 and 0.41.

In the present study the disease categories all diseases, udder diseases, and metabolic diseases were analysed with test day threshold models. Firstly, the trajectories of the different disease categories are described. Afterwards, variance components were estimated, where data recording comprises the first 50, 100, and 300 days of lactation, respectively. Furthermore, the effect of an increasing number of daughters per sire (improved genetic structure of the data) was analysed.

Materials and methods

Data recording took place from February 1998 to December 2002 on three commercial milk farms with an overall total of 3200 German Holstein cows. A special data recording scheme for functional traits (e.g. health traits) was established on these farms. In this data recording scheme the information, which was stored on farms in computer based management programs is used for a further improvement of the performance test in dairy cattle. Therefore farms were visited monthly and security copies (Backups) of the computer based management programs were taken. New information from the previous month was checked for plausibility. Questionable new information was discussed with the farm management and the veterinarian and discarded if they could not be clearly allocated. Furthermore, the same bulls were used for artificial insemination on the three farms. This led to a genetic connection between

the farms. Disease information was recorded by veterinarian or farm staff. During the first 300 days of lactation 75,595 disease treatments were recorded in total and thereof 39,934 come under udder disease treatments (Table 1).

Table 1

Number of recorded disease treatments for the first 50, 100, and 300 days in milk (Anzahl der Behandlungen für die ersten 50, 100 und 300 Laktationstage)

Disease category		Days in milk		
		50	100	300
Udder diseases	UD	17,122	23,209	39,934
Fertility diseases	FD	16,203	19,638	24,563
Metabolic diseases	MD	3410	3615	4131
Claw and leg diseases	CD	2138	2916	5085
All diseases except Fertility diseases	ADEFD	23,751	31,029	51,032
All diseases	AD	39,954	50,667	75,595

Table 1 summarises the recorded treatments of the analysed disease categories during the first 50, 100, and 300 days of lactation. The disease category ‘all diseases except fertility diseases’ was created for analysing the impact of fertility diseases on the results of the variance component estimation for all diseases. In general, the data sets could be divided according to lactation length and the number of daughters per sire (Table 2 and Table 3). Table 2 gives a description of the data sets 1 to 9, whereas Table 3 describes the data sets 10 to 18.

Table 2

Number of observations, cows, and sires in data sets 1 – 9, with different disease categories (udder, metabolic, and all diseases) and length of lactations (50, 100, and 300 days) (Anzahl der Beobachtungen je Krankheitskategorie und Anzahl der Kühe und Bullen in den Datensätzen 1 bis 9)

DS	DC	LL	Observations	Cows	Sires
1	UD	50 days	839,473	10,071	767
2	UD	100 days	1,612,471	10,210	783
3	UD	300 days	4,361,909	10,521	803
4	MD	50 days	751,462	9968	763
5	MD	100 days	1,440,692	10,317	780
6	MD	300 days	3,908,830	10,456	801
7	AD	50 days	839,473	10,071	767
8	AD	100 days	1,612,471	10,210	783
9	AD	300 days	4,361,909	10,521	803

DS = data set, DC = disease category, UD = udder diseases, MD = metabolic diseases, AD = all diseases, LL = lactation length
All available information used

Table 3

Number of observations, cows, and sires in data sets 10 – 18, with different disease categories (udder, metabolic, and all diseases) and restriction for minimum number of daughters per sire (at least 10, 30, or 50 daughters) (lactation length 50 days) (Anzahl der Beobachtungen, Kühe und Bullen für die verschiedenen Krankheitskategorien, sowie die minimale Anzahl der Töchter pro Bulle)

DS	DC	NOD	Observations	Cows	Sires
10	UD	at least 10	691,067	8356	209
11	UD	at least 30	519,855	6406	87
12	UD	at least 50	436,751	5482	61
13	MD	at least 10	549,452	7124	204
14	MD	at least 30	399,147	5266	84
15	MD	at least 50	333,705	4482	60
16	ADEFD	at least 10	691,067	8356	209
17	ADEFD	at least 30	519,855	6406	87
18	ADEFD	at least 50	436,751	5482	61

DS = data set, DC = disease category, UD = udder diseases, MD = metabolic diseases, ADEFD = all diseases except fertility diseases,
NOD = number of daughters; Data recording concentrated on the first 50 days of lactation

The effect of an increasing number of daughters per sire was analysed in the data sets 10 to 18. This was equal to an improved genetic structure of the data. This improved genetic structure of the data was caused by the special data recording scheme.

In all data sets disease information was treated as an ‘all or none’ trait. All analysed data sets contained one observation per cow and day. Each observation (day) received a disease code, ‘1’ if a cow showed a disease and ‘0’ if not. If there was a recorded treatment, the day of treatment and the following five days are coded with ‘1’. The five day period was selected, because it is the average time in which milk has to be discarded. All other days were coded with ‘0’. In a first step all data sets were checked for extreme categories. These categories were excluded from the analysis, but it should be noted that this was not a problem in our data, and therefore nearly all data were included in the following analyses. The significance of fixed effects was analysed by using the GENMOD procedure of the SAS package (SAS, 1999). Statistical testing for each factor was done by removing this particular factor from the model and comparing the likelihood of this reduced model with the likelihood of the full model. The age of calving was in eight classes: (20 – 24), (25 – 29), (30 – 49) month in first parity, (30 – 39), (40 – 45), and (46 – 56) month in second parity, third parity (45 – 75) month, and higher than 3 (≥ 75 month). Furthermore, the model includes the herd calving season as a multi-code. Each year was divided into three calving seasons. Calving between January and March was summarised to calving season one, and April to August, and September to December are the other two calving seasons, respectively. A further systematic environmental effect was the herd*week (of observation) of the test day, where one herd*week (of observation) class was equal to one herd week. The influence of all systematic environmental effects were significant ($p < 0.001$). A threshold liability model (WRIGHT, 1934; DEMPSTER and LERNER, 1950; GIANOLA and FOULLEY, 1983) was used for the estimation of random permanent environmental variance and additive genetic variance. In this study an animal model was applied to all data sets.

The following test day repeatability threshold model was applied to the data sets 1 to 18:

$$E [\pi_{ijklm}] = \Phi (B_i + K_j + L_k + f_{k(\text{days in milk})} + m_l + t_m),$$

where:

$E [\pi_{ijklm}]$	= expected probability for occurrence of any disease
Φ	= cumulative probability function of the standard normal distribution
B_i	= fixed effect of the i - th herd*week (of observation) ($i = 1, \dots, 771$) *
K_j	= fixed effect of the j - th herd*year*season (of calving) ($j = 1, \dots, 45$)
L_k	= fixed effect of the k - th age of calving ($k = 1, \dots, 8$)
$f_{k(\text{days in milk})}$	= lactation curve nested within age of calving k
m_l	= random permanent environmental effect ($l=1, \dots, 10,071$) **
t_m	= random effect of the m - th animal ($m=1, \dots, 15,802$)

* The number of herd*year*season effects (of observation) decreased for metabolic diseases from 771 to 421. This was caused by extreme categories, which were not included into the analyses.

** The number of random permanent environmental effects was 10,071 for data set 1. For the remaining data sets the number of random permanent environmental effects are given in Table 2 and Table 3. A pedigree file was applied to all data sets, which contains as much information as possible from two ancestor generations.

In the test day models for udder diseases, all diseases, and ‘all diseases except fertility diseases’ $f_{k(\text{days in milk})}$ represents a fixed effect with seven classes for the first seven days of lactation. After the seventh day of lactation $f_{k(\text{days in milk})}$ is identical to the function used by ALI and SCHAEFFER (1987). For metabolic diseases $f_{k(\text{days in milk})}$ represents a fixed effect with 18 classes for the first 18 days of lactation. After the 18th day of lactation $f_{k(\text{days in milk})}$ is identical to the function used by ALI and SCHAEFFER (1987).

The posterior distributions of the permanent environmental variance and the additive genetic variance for the liability to udder diseases or all diseases were determined with the Gibbs sampling algorithm implemented in the LMMG_TH program, a threshold derivative of LMMG (REINSCH, 1996). The LMMG_TH program based on the results presented by SÖRENSEN et al. (1995). For random effects multivariate normal distribution with zero means and appropriate variance – covariance matrices were used and improper flat priors for fixed effects. For all models 100,000 cycles were generated and the result from each cycle was retained. The results of the first 10,000 cycles were discarded (burn in plus a safety margin) and the results of the remaining 90,000 cycles were used to calculate the genetic parameters, using the MEAN procedure of the SAS package (SAS, 1999). The convergence was determined by visual inspection of plots of realised parameter values against iteration number. Similar convergence detection was used by REINSCH et al. (1999).

Results

Trajectories of different disease categories

The trajectory of udder diseases showed only one peak during the first 300 days of lactation (Figure 1). Thereafter the frequency of udder diseases decreased up to a constant level of 2% to 3%.

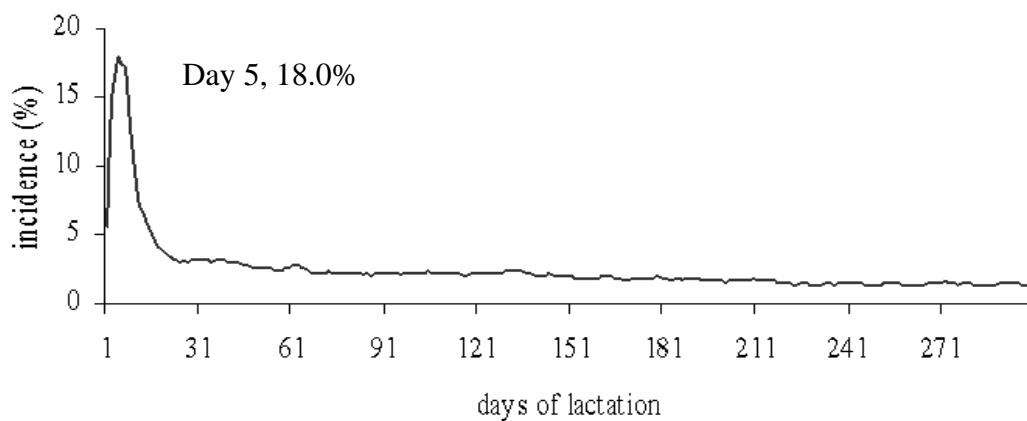


Fig. 1: Frequency of udder diseases during the first 300 days of lactation (Frequenz der Eutererkrankungen während der ersten 300 Laktationstage)

The disease incidence of metabolic diseases showed a further peak at lactation day 14 (Figure 2). Thereafter the metabolic disease incidence decreased and was below 1% after the 24th day of lactation.

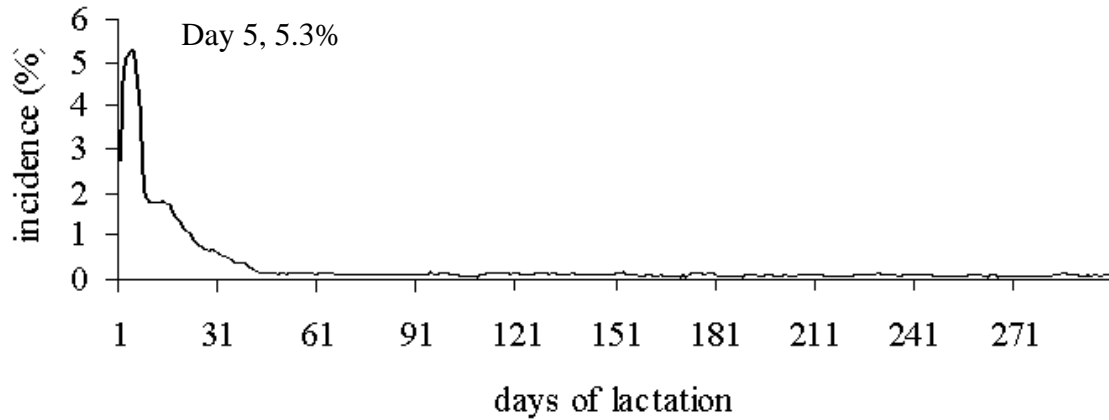


Fig. 2: Frequency of metabolic diseases during the first 300 days of lactation (Frequenz der Stoffwechselerkrankungen während der ersten 300 Laktationstage)

Figure 3 describes the trajectory of the disease category all diseases.

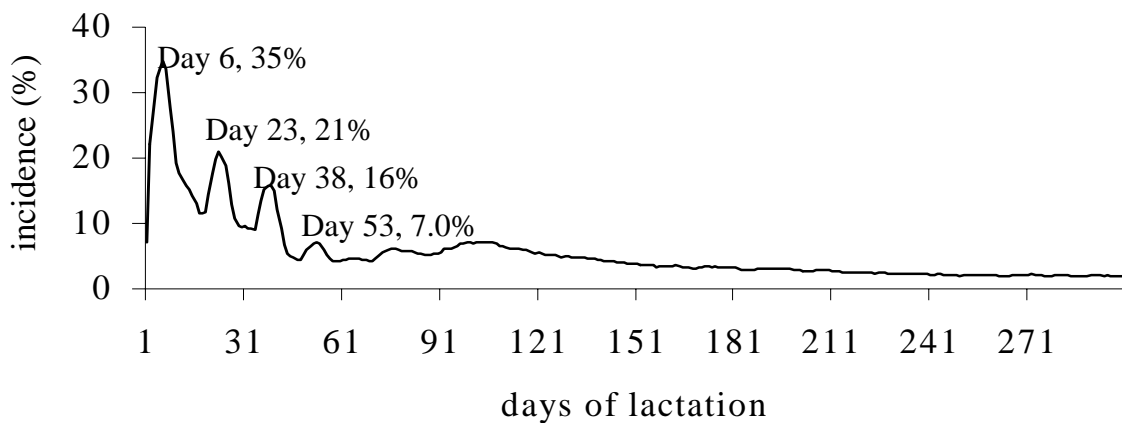


Fig. 3: Frequency of all diseases during the first 300 days of lactation (Frequenz aller Erkrankungen während der ersten 300 Laktationstage)

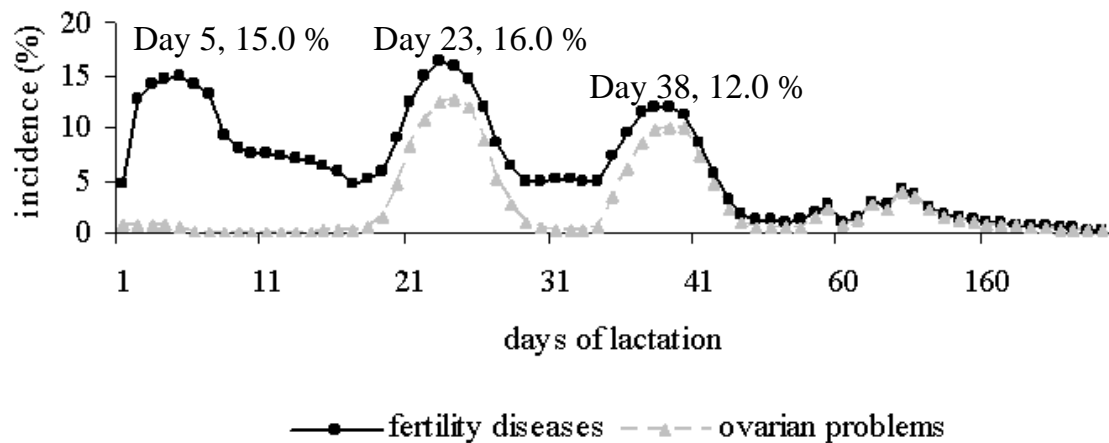


Fig. 4: Frequency of fertility diseases and ovarian problems (Frequenz von Fruchtbarkeitserkrankungen und Zyklusstörungen)

Figure 1 and Figure 3 suggest that the first peak of the trajectory of all diseases was caused by udder diseases, whereas the distinct peaks in the latter part of lactation must be caused by other health problems. The high number of fertility disease treatments (Table 1) pointed out, that fertility diseases could be the reason for these peaks.

The trajectories of fertility diseases and ovarian problems are shown in Figure 4. The trajectory of fertility diseases showed three distinct peaks, where the first peak was caused by non-ovarian problems, and the other two were caused by ovarian problems. All in all, the Figures 1 to 4 described the trajectories of the most important disease categories. The first peak of the trajectory of all diseases (Figure 3) summarises the non-ovarian problem peak (Figure 4) and the peak of the udder health problems and metabolic problems (Figure 1 and Figure 2). The remaining peaks in Figure 3 were caused by an increased risk of ovarian problems (Figure 4). In addition to the peaks the trajectory of all diseases was not so constant as the trajectory of udder diseases. Fertility diseases were excluded for the estimation of the influence of an increasing number of daughters per sire. This was done because the distinct peaks of the trajectory could create problems if variance components were to be estimated.

Disease frequencies

Table 4 summarises the average incidence of the disease categories in the data sets 1 to 9. As expected the disease category all diseases showed the highest frequencies, followed by udder diseases and metabolic diseases.

Table 4

Frequencies of udder diseases, metabolic diseases, and all diseases depending on lactation length (Erkrankungsfrequenzen für Euter-, Stoffwechsel- und alle Erkrankungen und Datenerfassungszeitraum für die Datensätze 1 bis 9)

DS	DC	LL	Frequency (%)
1	Udder diseases	50 days	5.9
2	Udder diseases	100 days	4.2
3	Udder diseases	300 days	2.7
4	Metabolic diseases	50 days	1.6
5	Metabolic diseases	100 days	0.9
6	Metabolic diseases	300 days	0.4
7	All diseases	50 days	15.3
8	All diseases	100 days	10.6
9	All diseases	300 days	6.0

DS = data set, DC = disease category, LL = lactation length

Table 5

Number of lactations (n), mean lactation length and standard deviation (SD) and frequencies (%) of the different disease categories for the first 300 days of lactation (all available information used) (Anzahl der Laktationen (n), mittlere Laktationslänge und Standardabweichung und Erkrankungsfrequenzen in den verschiedenen Krankheitskategorien innerhalb der ersten 300 Laktationstage)

Lactation	n	Mean LL (SD)	UD	MD	ADEFD	AD
1	8031	143 (87)	2.1	0.2	3.8	5.3
2	5256	142 (85)	2.7	0.3	4.2	5.6
3	3628	142 (86)	3.1	0.5	5.1	6.6
> 3	4453	138 (85)	3.3	0.7	6.0	7.5
Total	21,368	142 (86)	2.7	0.4	4.6	6.0

LL = lactation length, UD = udder diseases, MD = metabolic diseases, ADEFD = all diseases except fertility diseases, AD = all diseases

For all analysed disease categories the disease frequency decreased with increasing lactation length. This suggests that disease problems are concentrated at the beginning of the lactation. Therefore the analysis of the effect of an increasing number of

daughters per sire was restricted to the first 50 days of lactation. Furthermore, fertility diseases were excluded from all diseases.

The disease frequencies of the different disease categories were fairly constant in data sets 10 to 18 and they were between 6.1% and 6.5%, 1.6% and 1.7%, and 8.1% and 8.2% for udder diseases, metabolic diseases, and all diseases except fertility diseases, respectively.

In Table 5 the disease frequencies of the different disease categories are given for the different lactations. Furthermore, this table includes the number of lactations and the average lactation length and its standard deviation.

Estimation of variance components

The results of the variance component estimation for the first 9 data sets are summarised in Table 6 and all estimates are posterior means.

Table 6

Posterior means and standard deviations (SD) of additive genetic variance (σ_a^2), permanent environmental variance (σ_{pe}^2), heritabilities (h^2), and repeatabilities (r) for liability to udder diseases, metabolic diseases, and all diseases in data sets including 50, 100, or 300 days of lactation (Posteriori Mittelwerte der additiv genetischen Varianz (σ_a^2), der permanenten Umweltvarianz (σ_{pe}^2), der Heritabilität (h^2) und der Wiederholbarkeit (r) der Datensätze 1 bis 9 und deren Standardabweichung)

DS	DC	LL	σ_a^2 (SD)	σ_{pe}^2 (SD)	h^2 (SD)	r (SD)
1	UD	50 days	0.58 (.07)	1.80 (.07)	0.17 (.02)	0.70 (.01)
2	UD	100 days	0.44 (.04)	1.30 (.05)	0.16 (.02)	0.64 (.01)
3	UD	300 days	0.24 (.03)	0.83 (.02)	0.12 (.01)	0.52 (.01)
4	MD	50 days	0.91 (.14)	3.16 (.15)	0.18 (.03)	0.80 (.01)
5	MD	100 days	0.78 (.13)	2.84 (.15)	0.17 (.03)	0.78 (.01)
6	MD	300 days	0.43 (.07)	2.12 (.08)	0.12 (.02)	0.72 (.01)
7	AD	50 days	0.07 (.01)	0.60 (.02)	0.04 (.01)	0.40 (.01)
8	AD	100 days	0.05 (.01)	0.40 (.01)	0.03 (.01)	0.31 (.01)
9	AD	300 days	0.04 (.01)	0.28 (.01)	0.03 (.01)	0.25 (.01)

DS = data set, DC = disease category, UD = udder diseases, MD = metabolic diseases, AD = all diseases, LL = lactation length

The estimated additive genetic variances vary between 0.04 (data set 9) and 0.91 (data set 4) and were between 0.28 and 3.16 for the permanent environmental variance, respectively. Furthermore, for all disease categories a decreasing additive genetic variance and permanent environmental variance could be observed with increasing lactation length (Table 6). The highest additive genetic variance (0.43 to 0.91), as well as permanent environmental variance (2.12 to 3.16) were estimated for the liability to metabolic diseases followed by the liability to udder diseases. The additive genetic variance of the liability to udder diseases were estimated within the interval of 0.24 to 0.58 and were between 0.83 and 1.80 for the permanent environmental variance.

The estimated variance components for the liability to all diseases were significantly lower, as compared to the liability to metabolic diseases or udder diseases. This led to lower heritabilities of liability (0.04 to 0.07) for all diseases, whereas similar heritabilities of liability were estimated for metabolic diseases (0.12 to 0.18) and udder diseases (0.12 to 0.17).

The effect of an increasing number of daughters per sire (improved genetic structure) on the variance component estimation is shown in Table 7. All estimates are posterior means.

The estimated additive genetic variances vary between 0.38 and 1.47, depending on the disease category and they were between 1.18 and 3.72 for the permanent environmental variance, respectively. Within all disease categories additive genetic

variance increased with an improved genetic structure of the data. The permanent environmental variance increases for the liability to metabolic diseases and ‘all diseases except fertility diseases’ with an increasing number of daughters per sire (improved genetic structure), whereas the permanent environmental variance decreased for the liability to udder diseases, respectively (Table 7).

Table 7

Posterior means and standard deviations (SD) of additive genetic variance (σ_a^2), permanent environmental variance (σ_{pe}^2), heritabilities (h^2), and repeatabilities (r), for liability to udder diseases, metabolic diseases, and all diseases except fertility diseases, in data sets with an improved genetic structure (different restrictions for minimum number of daughters per sire, at least 10, 30, or 50 daughters) (Posteriori Mittelwerte der additiv genetischen Varianz (σ_a^2), der permanenten Umweltvarianz (σ_{pe}^2), der Heritabilität (h^2) und Wiederholbarkeit der Datensätze 10 bis 18 (verbesserte genetische Struktur) und deren Standardabweichungen)

DS	DC	NOD	σ_a^2 (SD)	σ_{pe}^2 (SD)	h^2 (SD)	r (SD)
10	UD	at least 10	0.63 (.07)	1.80 (.08)	0.18 (.02)	0.71 (.01)
11	UD	at least 30	0.74 (.09)	1.70 (.09)	0.22 (.02)	0.71 (.01)
12	UD	at least 50	0.90 (.14)	1.68 (.11)	0.25 (.03)	0.72 (.01)
13	MD	at least 10	0.78 (.15)	3.00 (.18)	0.16 (.03)	0.79 (.01)
14	MD	at least 30	1.01 (.15)	3.28 (.23)	0.19 (.03)	0.81 (.01)
15	MD	at least 50	1.47 (.30)	3.72 (.33)	0.24 (.04)	0.84 (.01)
16	ADEFD	at least 10	0.38 (.04)	1.18 (.05)	0.15 (.01)	0.61 (.01)
17	ADEFD	at least 30	0.50 (.06)	1.22 (.07)	0.18 (.02)	0.63 (.01)
18	ADEFD	at least 50	0.53 (.08)	1.27 (.08)	0.19 (.03)	0.64 (.01)

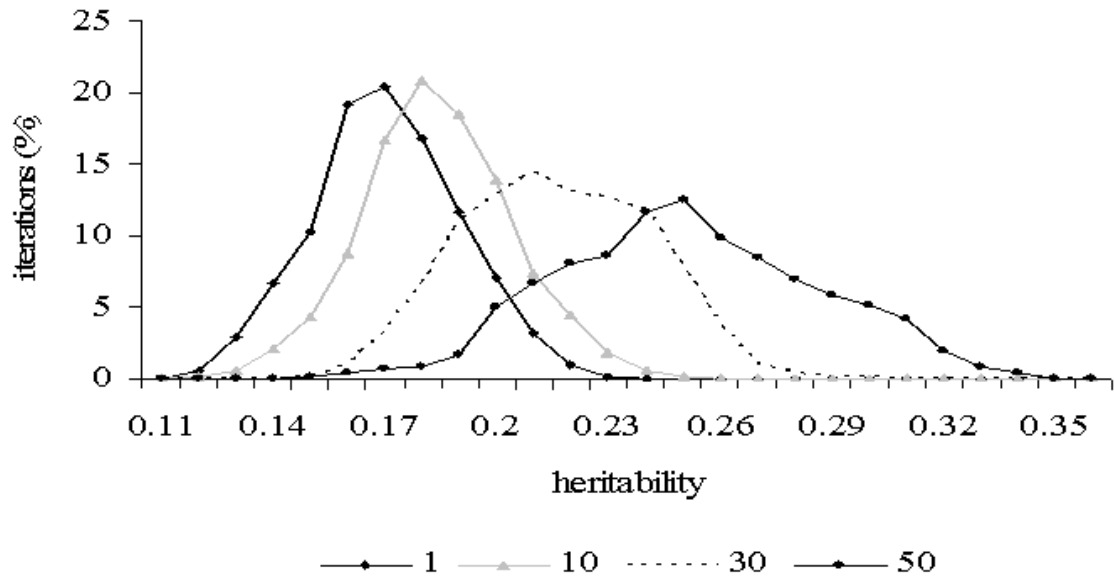
DS = data set, DC = disease category, UD = udder diseases, MD = metabolic diseases, ADEFD = all diseases except fertility diseases, NOD = number of daughters

The estimated heritabilities of liability were within the interval of 0.15 to 0.25. With an increasing number of daughters per sire (improved genetic structure) an increasing heritability of liability could be observed for all analysed disease categories. The increasing heritability of liability was more distinct for the disease categories udder diseases (0.18 to 0.25) and metabolic diseases (0.16 to 0.24) compared to the disease category ‘all diseases except fertility diseases’ (0.15 to 0.19). The increasing heritability of liability showed that differences between sires become more apparent with increasing number of daughters per sire (improved genetic structure). Furthermore, the genetic connection between the farms should have a positive effect on the heritabilities.

For all analysed data sets the estimated repeatabilities were high with values within the interval of 0.61 to 0.84 (Table 7).

Posterior distributions are shown exemplary for the liability to udder diseases in Figure 5, where the considered period of lactation ends after 50 days in milk. Posterior distributions of the heritability were similar for the data sets, which contain all available udder disease information, as well as for the data set when only records from sires with at least 10 daughters were considered. Furthermore, posterior distributions of the heritability of the two remaining data sets were similar, but they are not so peaked (Figure 5).

The posterior distributions for the liability to metabolic diseases (not shown) were more flat, compared to those of the liability to udder disease. In addition the posterior distributions of the liability to all diseases except fertility diseases (not shown) were similar to those shown in Figure 5.



1 = at least 1 daughter, 10 = at least 10 daughters, 30 = at least 30 daughters, 50 = at least 50 daughters

Fig. 5: Posterior distribution of heritability of the liability for udder diseases (first 50 days of lactation) for data sets with an improved genetic structure (at least 1, 10, 30, or 50 daughters per sire) (Posteriori Verteilungen der Heritabilität von Eutererkrankungen für die ersten 50 Laktationstage bei verbesserter genetischer Struktur)

Discussion

The aggregation of different diagnoses to disease categories was very useful for this analysis. On the one hand inconsistent diagnosis given by the veterinarian could be avoided. On the other hand the definition of disease categories reduced the number of discarded information. For example, disease categories reduced the number of extreme categories, because similar diagnoses were summarised.

If test day models are used the trajectory of the analysed trait has an important effect. Therefore, models must be fitted to the trajectory of the analysed disease category. For udder diseases, metabolic diseases, and 'all diseases except fertility diseases' this could be done by using the Ali Schaeffer function, but it should be noted that the possibilities of these function are also limited. Therefore, the first days of lactation must be treated as fixed effects. This was also done by SCHOMAKER et al. (2002).

Disease frequencies and estimation of variance components

The range of frequencies of different disease categories or diseases vary widely in the literature, depending on data recording methods and case definition. Furthermore, the frequencies of diseases categories are affected by the number of diseases, which were summarised to one disease category. NIELSEN et al. (1999) decided between udder diseases, reproductive diseases, digestive diseases, and feet and leg diseases. Each of these disease categories summarises nearly ten different diagnoses. Other studies (LUND et al., 1994 and HANSEN et al., 2002) divided disease information into two disease categories, mastitis and diseases other than mastitis. There is no standardisation for the creation of disease categories. Standardisation for the evaluation of disease categories would improve the comparability of different analyses.

Overall, in the analysed disease categories the average disease frequencies decreased with increasing lactation length (Table 4). This suggests that all disease problems are

concentrated in a relatively short phase at the beginning of lactation, when the physiological demands of the cows are high. For udder diseases the concentration of diseases at the beginning of lactation is in agreement with the results found by HERINGSTAD et al. (2003b), KADARMIDEEN et al. (2000), and REKAYA et al. (1998). An increasing number of daughters per sire (improved genetic structure) showed no effect on the average disease frequency. It should be noted that in this study all lactations were included, whereas most other studies based on first lactation information.

For udder diseases the additive genetic variance and the permanent environmental variance decreased with the increasing period of data collection (Table 6). This resulted in decreasing heritabilities (0.17 to 0.12) and repeatabilities (0.70 to 0.52). These estimates are higher than those found by REKAYA et al. (1998), KADARMIDEEN et al. (2000), and HERINGSTAD et al. (2003b) for mastitis if test day threshold models were used. These authors divided the lactation in several blocks. This could be a reason for the different results. Another possibility for the different results could be that the authors mentioned above used sire threshold models. In a previous analysis (results not shown) results were checked by a second analysis with a sire model. In our case the results of both models agreed well. Additionally, the estimated heritabilities for udder diseases in this study are significantly higher than those given by several authors (SIMIANER et al., 1991; HERINGSTAD et al., 1997; HERINGSTAD et al., 2003a) who carried out their analyses with lactation threshold models.

The additive genetic variance and the permanent environmental variance of metabolic diseases also decreased with increasing lactation length. The additive genetic variance and the permanent environmental variance of metabolic diseases were higher than those estimated for udder diseases and all diseases. However, the estimated heritabilities (0.18 to 0.12) are similar to the heritabilities of udder diseases and they are in the upper range of literature reports (SIMIANER et al., 1991; LYONS et al., 1991; URIBE et al., 1995). It should be noted that in most cases cows are only affected by one metabolic disease during lactation. Therefore, the improvement in genetic parameter estimation, caused by the use of test day threshold models is limited. This is not correct for udder diseases where repeated cases occurred in a higher frequency, compared to metabolic diseases. This resulted in an improvement in genetic parameter estimation if test day threshold models are used for genetic analysis of udder disease data because distinct cases of udder diseases could be treated as distinct observations in those models (HINRICHS et al., 2003).

If all available disease information was analysed simultaneously the additive genetic variance (0.04 to 0.07), as well as the permanent environmental variance (0.28 to 0.40) are low, compared to the disease categories udder diseases and metabolic diseases.

The use of lactation threshold models resulted in low estimates for the additive genetic variance and the permanent environmental variance. In addition, the estimated heritabilities and repeatabilities are low (HINRICHS et al., 2004). If test day threshold models were used the additive genetic variance was fairly similar compared to the additive genetic variance of the lactation threshold models. Therefore, heritabilities did not increase with the use of test day threshold models in contrast to the repeatabilities (Table 6). The estimated heritabilities are similar to those estimated by SIMIANER et al. (1991). The fairly constant additive genetic variance could be caused by the

trajectory of all diseases (Figure 3). These trajectories showed distinct peaks and some of them were caused by ovarian problems (Figure 4). Therefore, ovarian problems cannot be analysed together with other diseases and should be assessed separately. An alternative function must be found for modelling the trajectory of ovarian problems, or the Ali Schaeffer function should be replaced by a fixed effect for every day. Such data sets have to be very large, because several observations will be needed for an exact estimation of the fixed effect of the lactation day.

The effect of an increasing number of daughters per sire (improved genetic structure) on the results of the variance component estimation was analysed with the disease categories udder diseases and metabolic diseases. In addition, fertility diseases were excluded from the disease category all diseases before the variance components were estimated. The lactation length was restricted to the first 50 days of lactation, because most diseases occurred during this period of lactation (Figures 1 to 4). An increasing number of daughters per sire (improved genetic structure) showed a positive overall effect in the analysed disease categories (Table 7). The results of the variance component estimation for the diseases category ‘all diseases except fertility diseases’ suggests that fertility diseases should be analysed separately (Table 6 and Table 7).

All in all the results of our study showed that an increasing number of daughters per sire (improved genetic structure) had a positive effect on the variance component estimation, if test day threshold models were used. This positive effect cannot be observed if the analyses were carried out with lactation threshold models (HINRICHS et al., 2004). Therefore, the development of test day threshold models is an important area of research. Alternative models should be tested for their possibilities to describe the trajectories of the disease categories all diseases and fertility diseases.

Another possibility could be the use of random regression models or multiple trait models. For such models very large data sets will be needed which are not available in most countries today. An advantage of random regression models would be that they treat diseases at different lactation days as different traits. This could be useful because HERINGSTAD et al. (2003b) found that the genetic correlation between mastitis in different stages of lactation was not equal to one. Therefore mastitis in different stages of lactation is not the same trait and should be treated as different traits, if it is possible.

Special data recording scheme

The results of our study showed that the establishment of a special data recording scheme for functional traits on commercial dairy farms is possible. Data recording on commercial dairy farms would lead to large field data sets, which are not available in most countries today. These data sets could be used for the estimation of genetic parameters and breeding values. The establishment of such data recording schemes should be done stepwise. In a first step potential data recording farms have to be selected. These farms have to fulfil several requirements. They should be as large as possible and farm management, farm staff, and veterinarians must be interested in the data recording scheme. Furthermore, all farms must have computer based herd management programs. All information should be stored on the farm as long as possible. The second step includes the construction of a data base for the data recording scheme and first data recording. Data recording has to become a part of the daily farm work and plausibility checks for the recorded information must be established. Thereafter the same bulls should be used for artificial insemination at the

same time on the farms. A genetic connection between the data recording farms would therefore be built up. If the data recording is well established on the farms another important point is, that the motivation of farm staff, farm management, and veterinarians for the data recording has to be maintained. This could be done in the form of disease reports or warnings about certain disease outbreaks.

Conclusions

The presented results suggest that disease traits showed sufficient additive genetic variance. Therefore, disease information should be used in commercial breeding schemes. Furthermore, disease information should be divided into disease categories. The development of test day threshold models is an important area of research. The trajectory of all diseases showed more peaks and could not be described by the Ali Schaeffer function. Therefore alternative functions should be tested for this disease category. In particular fertility diseases should be analysed separately. In addition, fertility diseases should be divided in ovarian and non ovarian problems. All models must be fitted to the trajectory of the disease category. For the genetic improvement of disease resistance data recording, quantitative genetic information and molecular genetic information should be combined and integrated stepwise into commercial breeding schemes.

Acknowledgement

This project was supported by the Federal Ministry of Consumer Protection, Food and Agriculture, Berlin and NOG Nord Ost- Genetic GmbH & Co. KG, Verden. Special thanks to Jörn Bennewitz and Norbert Reinsch for helpful discussion.

References

- ALI, T.E; SCHAEFFER, L.R.:
Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* **67** (1987), 637–644
- COLLARD, B.L.; BOETTCHER, P.J.; DEKKERS, J.C.M.; PETITCLERCT, D.; SCHAEFFER, L.R.:
Relationships between energy balance and health traits of dairy cattle in early lactation. *J. Dairy Sci.* **83** (2000), 2683-2690
- DEMPSTER, E.R.; LERNER, I.M.:
Heritability of threshold characters. *Genetics* **35** (1950), 212–235
- GIANOLA, D.; FOULLEY, J.L.:
Sire evaluation for ordered categorical data with a threshold model. *Genet. Sel. Evol.* **15** (1983), 201–224
- HANSEN, M.S.; LUND, K.M.; SÖRENSEN, K.; CHRISTENSEN, L.G.:
Genetic parameters of dairy character, protein yield, clinical mastitis, and other diseases in Danish Holstein cattle. *J. Dairy Sci.* **85** (2002), 445-452
- HERINGSTAD, B.; KARLSEN, A.; KLEMETSDAL, G.; RUANE, J.:
Preliminary results from a genetic analysis of clinical mastitis data. *Interbull Bull.* **15** (1997), 45 – 49
- HERINGSTAD, B.; KLEMETSDAL, G.; RUANE, J.:
Selection for mastitis resistance in dairy cattle—A review with focus on the situation in the Nordic countries. *Livest. Prod. Sci.* **64** (2000), 95–106
- HERINGSTAD, B.; REKAYA, R.; GIANOLA, D.; KLEMETSDAL, G.; WEIGEL, K.A.:
Bivariate analysis of liability to clinical mastitis and to culling in first-lactation cows. *J. Dairy Sci.* **86** (2003a), 653–660
- HERINGSTAD, B.; CHANG, Y.M.; GIANOLA, D.; KLEMETSDAL, G.:
Genetic analysis of longitudinal trajectory of clinical mastitis in first-lactation Norwegian Cattle. *J. Dairy Sci.* **86** (2003b), 2676–2683
- HINRICH, D.; STAMER, E.; JUNGE, W.; KALM, E.:

- A comparison of models for genetic evaluation of mastitis data: Heritabilities, repeatabilities and genetic correlation with milk production traits. Book of Abstracts page 82. 54th Annual Meeting of EAAP, Rome, Italy (2003)
- HINRICHS, D.; STAMER, E.; JUNGE, W.; KALM, E.:
Genetic analysis of several economically important disease traits in German Holstein cows. Submitted for publication (2004).
- JAKOB, H.; DISTL, O.:
Tierarztkosten beim Milchvieh. 2 Mitteilung: Beziehungen zwischen Tierarztkosten und der Milchleistung beim Deutschen Fleckvieh und Deutschem Braunvieh. *Züchtungskunde* **70** (1998), 29 – 42
- KADARMIDEEN, H.N.; REKAYA, R.; GIANOLA, D.:
Genetic analysis of repeated mastitis episodes with Bayesian longitudinal threshold models. Book of Abstracts page 78. 51th Annual Meeting of EAAP, The Hague, The Netherlands (2000)
- KELTON, D.F.; LISSEMORE, K.D.; MARTIN, R.E.:
Recommendations for recording and calculating the incidence of selected clinical diseases of dairy cattle. *J. Dairy Sci.* **81** (1998), 2502-2509
- LIN, H.K.; OLTENCAU, P.A.; VAN VLECK, L.D.; ERB H.N.; SMITH, R.D.:
Heritabilities of and genetic correlations among six health problems in Holstein cows. *J. Dairy Sci.* **72** (1989), 180-186
- LUND, T.; MIGLIOR, F.; DEKKERS, J.C.M.; BURNSIDE, E.B.:
Genetic relationship between clinical mastitis; somatic cell count, and udder conformation in Danish Holsteins. *Livest. Prod. Sci.* **39** (1994), 243–251
- LYONS, D.T.; FREEMAN, A.E.; KUCK, A.L.:
Genetics of health traits in Holstein cattle. *J. Dairy Sci.* **74** (1991), 1092–1100
- NIELSEN, U.S.; PEDERSEN, G.A.; PEDERSEN, J.; JENSEN, J.:
Genetic variation in disease traits and their relationships with survival in Danish dairy Cattle. *Interbull Bull.* **21** (1999), 170–177
- REINSCH, N.:
Two Fortran programs for the Gibbs Sampler in univariate linear mixed models. *Arch. Tierz., Dummerstorf* **39** (1996), 203–209
- REINSCH, N.; ENGELLANDT, T.; SCHILD, H.- J.; KALM, E.:
Lack of evidence for bovine Y-chromosomal variation in beef traits. A Bayesian analysis of Simmental data. *J. Anim. Breed. Genet.* **116** (1999), 437 – 445
- REKAYA, R.; RODRIGUEZ-ZAS, S.L.; GIANOLA, D.; SHOOK, G.E.:
Test-day models for longitudinal binary response: an application to mastitis in Holsteins. Book of Abstracts page 44. 49th Annual Meeting of EAAP, Warsaw, Poland (1998)
- SAS, 1999:
SAS/STAT User's guide, Version 8, SAS Institute, Cary, NC, USA
- SCHOMAKER, T.; JUNGE, W.; STAMER, E.; REINSCH, N.; KALM, E.:
Varianzkomponentenschätzung für die Merkmale Mastitis und Zellgehalt (SCS) und Vergleich verschiedener Auswertungsmodelle. *Züchtungskunde* **74** (2002), 250 – 265
- SIMIANER, H.; SOLBU, H.; SCHAEFFER, L.R.:
Estimated genetic correlations between disease and yield traits in dairy cattle. *J. Dairy Sci.* **74** (1991), 4358–4365
- SÖRENSEN, D.; ANDERSEN, S.; GIANOLA, D.; KORSGAARD, I. R.:
Bayesian inference in threshold models using Gibbs sampling. *Gen. Sel. Evol.* **27** (1995), 229-249
- URIBE, H.A.; KENNEDY, W.; MARTIN, S.W.; KELTON, D.F.:
Genetic parameters for common health disorders of Holstein cows. *J. Dairy Sci.* **78** (1995), 421-430
- WRIGHT, S.:
An analysis of variability in number of digits in an inbred strain of Guinea pigs. *Genetics* **19** (1934), 506–536

Received: 2005-04-26

Accepted: 2005-12-12

Author's address

Dr. DIRK HINRICHS, Dr. ECKHARD STAMER, Dr. WOLFGANG JUNGE,

Prof. Dr. Dr. h.c. mult. ERNST KALM

Tierzucht und Tierhaltung, Christian-Albrechts-Universität zu Kiel

Herrmann-Rodewald-Strasse 6, 24118 KIEL, GERMANY

E-mail: dhinrichs@tierzucht.uni-kiel.de

E-mail: estamer@tierzucht.uni-kiel.de