Estimated Genetic Correlations Between Disease and Yield Traits in Dairy Cattle

H. SIMIANER
Department of Animal Husbandry and Animal Breeding
University of Hohenheim
PO Box 700562
D-7000 Stuttgart 70, Germany

H. SOLBU
Department of Animal Science
Agricultural University of Norway
N-1432 As-NLH, Norway

L. R. SCHAEFFER
Center for Genetic Improvement of Livestock
Department of Animal and Poultry Science
University of Guelph
Guelph, ON N1G 2W1, Canada

ABSTRACT

Data included observations on more than 200,000 first lactations of Norwegian cattle. Milk yield, fat and protein percentages, and observations on mastitis, ketosis, and presence of disease (binary coding of 0 or 1) were analyzed. Following Bayesian principles and applying the threshold concept, dispersion parameters for the binary traits (on the underlying scale) with continuous production traits were estimated. Heritabilities were .27, .34, and .43 for milk yield, fat percentage, and protein percentage, respectively. Heritabilities for the disease traits were .05 to .13 but may be inconsistent because of methodology problems with small sire by herd-year-season subclass size. Genetic correlations between milk yield and all three disease traits were above .5, indicating an undesirable relationship. Genetic correlations between ketosis and the content traits were −.38 to −.65; low component percentages were associated with high ketosis frequencies. Ignoring diseases in breeding programs may lead to undesirable correlated selection responses when selecting on milk yield.

(Key words: genetic correlation, disease traits, production traits)

INTRODUCTION

When selecting for milk yield, diseases should be kept at minimum levels from both economical and ethical points of view. The economic side of this is obvious. Estimation of exact economic losses caused by diseases is difficult and will differ widely from population to population. It has been shown how diseases and other traits can be accounted for in the estimation of the efficiency of different breeding programs (9).

The ethical aspects of disease monitoring are connected to animal welfare and consumer interests. In general, consumers want animal products produced by healthy animals with as little use of antibiotics and other drugs as possible.

In most conventional dairy cattle breeding programs, the selection pressure is mainly on milk yield. Susceptibility to diseases is included in the breeding objectives of some Scandinavian countries; for a review, see Emanuelson (1). The future development of disease frequencies in most populations will largely depend on the genetic correlations between susceptibility to these diseases and milk yield.

Literature on estimates of genetic correlations among milk yield and diseases was reviewed recently (1, 16). Although shortcomings of individual studies with respect to the methodology applied or the data used may be criticized, the results as a whole provide an adequate indication of the relationship of yield and diseases. Most problems arise from the
discrete nature of disease observations, which in most cases has not been taken into account. Therefore, estimates of the correlations depend on the disease frequency in the respective population and are expected to be biased towards 0. Emanuelson (1) advocates strongly the use of more appropriate statistical methods for all-or-none data in this context but adds that applications of these techniques in practical situations have been limited because of the complexity and high computer demand of such methods.

The objective of this study was to analyze the relationship of yield and diseases in a specific population. Appropriate statistical tools have been suggested (14) and were used in the present study for the first time to analyze data of considerable scope. Good estimates of the respective population parameters are essential for an adequate incorporation of both traits in breeding programs.

MATERIALS AND METHODS

First lactation records of Norwegian cattle from the years 1978 to 1983 were analyzed. The records were required to contain sire and herd identification, calving date, milk yield, and fat and protein percentages. Only lactations of at least 250 d were included in this study. Lactations longer than 250 d but shorter than 305 d were not extended except for cows that were culled in that period. The 305-d yield was estimated from monthly recordings and corrected for the length of the calving interval.

Data were restricted to completed first lactations. Including later lactations would call for a multitrait model to account for selection bias (1), which would further increase the computational difficulty. In addition, selection decisions are mainly based on first lactation results to keep generation intervals at a minimum. Therefore, the main interest is in parameters for the first lactation.

Diseases are likely to occur soon after calving and may result in involuntary early culling of cows. The culling decision depends on the (expected) yield of a cow: in the case of a disease, poor milkers will more likely be culled than good ones. Extension of short lactations to 305-d lactations seems problematic if a cow was culled at an early stage of lactation and only few test day records are available. In this case, a method should be used that allows for incomplete observations for the yield traits along with complete observations for diseases. This type of data cannot be handled with the methodology currently available (14).

In Norway, all veterinary treatments in milk recorded herds are recorded in the health card system (15). Because the use of drugs in veterinary medicine is heavily regulated, almost all cattle diseases in Norway are treated by a veterinarian. The disease recordings should therefore be quite complete.

Diseases were recorded in binary form, i.e., 0 for no observations of a particular disease and 1 for one or more observations of a disease in a given lactation. In addition to the two most frequent single disease events, mastitis and ketosis, a trait called "any disease" was considered, which is information on whether or not a cow was treated for any reason (including mastitis and ketosis) by a veterinarian during the first lactation. In total, 216,565 records met these criteria and were complete with respect to the yield and disease traits included.

From the sires represented in these records, 586 could be identified as test sires from the breeding program, which are expected to be mated to a random sample of cows. The test sires were a random effect in the model and had on average 200 daughters. Those 118 of the remaining sires with more than 100 daughters with records (1000 daughters on average) were considered as proven sires and were regarded as fixed to improve the structure of the data. Data from the daughters of the non-test sires with less than 100 daughters were discarded.

Using four seasons (January to March, April to June, July to September, and October to December), observations were assigned to herd-year-season subclasses. Each herd-year-season class was required to have daughters of at least two different test sires. Neighboring subclasses were joined until this criterion was met. Data from herds in which only one test sire had daughters over the whole period were eliminated.

Subsets of data had to be formed because of computational limitations, and each subset was edited as described previously. Initially, three subsets were formed according to years (1978 and 1979, 1980 and 1981, 1982 and 1983).
After analyzing the first subset, improvements in the computing strategy allowed combination of the second and third subset so that two subsets of different sizes were left. Computational limitations did not allow formation of new subsets of about equal size, which would have required reanalysis of one-third of the data. The numbers of observations, test and proven sires, and herd-year-seasons in each subset are given in Table 1.

The basic methodology for the estimation of covariance components between one continuous and one binary trait was described in detail by Simianer and Schaeffer (14). The main assumption is the presence of an underlying bivariate normal distribution, for the nonobservable liability to the disease being subdivided into two discrete observation classes by an abrupt threshold. Following Bayesian arguments, an algorithm was proposed to estimate location parameters (3) that has been generalized to estimate both location and dispersion parameters simultaneously (14).

Although a generalization of the methodology to several continuous or discrete traits is rather straightforward in principle, the amount of data and the complexity of the algorithm made it necessary to analyze only pairs of production and disease traits at one time. The statistical model for the underlying variables, observable yield trait ($i = 1$), and nonobservable susceptibility value ($i = 2$), was

$$y_{ijkl} = h_y s_{i1} + s_{j1} + e_{ijkl}$$

where

- $y_{ijkl}$ is the value of variable $i$ of daughter $k$ from sire $j$ in herd-year-season $i$;
- $h_y s_{i1}$ is the fixed effect of herd-year-season $i$ on variable $i$;
- $s_{j1}$ is the effect of sire $j$ on variable $i$ (fixed for the first $j_1$ sires and random for the sires $j_1 + 1$ to $j$); and
- $e_{ijkl}$ is the residual error term.

In matrix notation, this can be written

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_{11} \\ u_{12} \\ u_{21} \\ u_{22} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

where $y_1$, $\beta_1$, and $e_1$ are the vectors of variable values, fixed herd-year-season effects, and residuals for trait $i$; $u_{im}$ is the vector of fixed ($m = 1$) or random ($m = 2$) sire effects on trait $i$; $X$ is the design matrix for the fixed herd-year-season effects (identical for both traits because only complete pairs of observations were taken into account); and $Z_m$ is the design matrix for fixed or random sires, respectively.

Using

$$y' = \begin{bmatrix} y_1' \\ y_2' \end{bmatrix}, \beta' = \begin{bmatrix} \beta_1' \\ \beta_2' \end{bmatrix}$$

$$e' = \begin{bmatrix} e_1' \\ e_2' \end{bmatrix}, u'_m = \begin{bmatrix} u'_{1m} \\ u'_{2m} \end{bmatrix}$$

and

$$X^+ = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix}$$

the first and second moments of the random variables on the underlying scale are

$$E \begin{bmatrix} y \\ u_2 \\ e \end{bmatrix} = \begin{bmatrix} X^+ \beta + Z_1^+ u_1 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{Var} \begin{bmatrix} y \\ u_2 \\ e \end{bmatrix} = \begin{bmatrix} G \otimes Z_2^+ A_{22} Z_2^+ & + R \otimes I & G \otimes Z_2^+ A_{22} \otimes I \\ G \otimes A_{22} & 0 \\ \text{symmetric} & \otimes I \end{bmatrix}$$

where $G$ and $R$ are the $2 \times 2$ variance-covariance matrices pertaining to random sires.
and residuals, $\otimes$ is the direct or Kronecker product (13), and $A_{22}$ is the numerator relationship matrix of the random sires represented in the data. Because the algorithm used is based on a system of nonlinear equations similar to the mixed model equations (4), the inverse of $A_{22}$ is required. This was obtained as follows.

The pedigrees of the random test sires were traced back as far as possible following the male paths only because no information on maternal ancestors was available. Sires only appearing in the pedigree formed the new group of sires $u_0$ (these sires were partly identical with the sires represented in $u_1$). Then, the common inverse numerator relationship matrix for the sires in $u_0$ and $u_2$ was constructed directly following Henderson's rules (5):

$$A^{-1} = \begin{bmatrix} A^{00} & A^{02} \\ A^{02'} & A_{22} \end{bmatrix}.$$  

According to the rules of inversion of partitioned matrices (13), $A_{22}^{-1}$ is obtained by absorbing $A^{00}$ into $A_{22}$:

$$A_{22}^{-1} = A_{22} - A^{02'}(A^{00})^{-1}A^{02}.$$  

The major objective of this study was to estimate the unknown dispersion parameters on the underlying scale contained in $G$ and $R$, which can be written

$$G = \begin{bmatrix} \sigma^2_{s1} & \text{Cov}_s \\ \text{Cov}_s & \sigma^2_{s2} \end{bmatrix};$$

$$R = \begin{bmatrix} \sigma^2_{e1} & \text{Cov}_e \\ \text{Cov}_e & \sigma^2_{e2} \end{bmatrix};$$

where

- $\sigma^2_{si}$ is the sire ($x = s$) or residual ($x = e$) variance of trait $i$ and
- $\text{Cov}_x$ is the sire ($x = s$) or residual ($x = e$) covariance between the two traits.

Heritabilities and additive genetic correlations were computed from covariances of paternal half-sibs, implying that these covariances are only due to additive genetic effects. Because the underlying variable for the discrete trait is only conceptual, the value of $\sigma^2_{e2}$ can be chosen deliberately so that five unknown dispersion parameters remain to be estimated. The following set of parameters describing completely the dispersion structure was chosen: heritability for the continuous trait,

$$h_c^2 = \frac{4\sigma^2_{s1}/(\sigma^2_{s1} + \sigma^2_{e1})}{2};$$

heritability for the disease trait,

$$h_d^2 = \frac{4\sigma^2_{s2}/(\sigma^2_{s2} + \sigma^2_{e2})}{2};$$

additive genetic correlation,

$$r_a = \frac{\text{Cov}_a}{(\sigma^2_{s1}\sigma^2_{s2})};$$

residual correlation,

$$r_e = \frac{(\text{Cov}_e - 3\text{Cov}_s)/(\sigma^2_{e1} - 3\sigma^2_{s1})}{(\sigma^2_{e1} - 3\sigma^2_{s1})}5;$$

and the residual variance for the continuous trait,

$$\sigma^2_e = \sigma^2_{e1} - 3\sigma^2_{s1}.$$  

Two alternative strategies to estimate $r_e$ were proposed by Simianer and Schaeffer (14). The first is an explicit search for the value of $r_e$ that maximizes the log posterior density of the location parameters given the observations and the dispersion parameters. An alternative is to use an approximate maximum likelihood approach (17), assuming multivariate normality and following the threshold concept. Because the latter proved to be computationally advantageous and was found empirically to lead to the same results as the former, this second computing strategy as described by Simianer and Schaeffer (14) was adopted for the present study.

The analysis results are given for each subset separately and also pooled in the form of...
TABLE 1. Numbers of observations, test sires, proven sires, and herd-year-season subclasses by data set.

<table>
<thead>
<tr>
<th></th>
<th>All data</th>
<th>Subset 1</th>
<th>Subset 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of records</td>
<td>208,693</td>
<td>54,118</td>
<td>138,418</td>
</tr>
<tr>
<td>Number of test sires</td>
<td>586</td>
<td>228</td>
<td>358</td>
</tr>
<tr>
<td>Number of proven</td>
<td>118</td>
<td>91</td>
<td>98</td>
</tr>
<tr>
<td>sires</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of herd-year-</td>
<td>71,406</td>
<td>19,202</td>
<td>46,481</td>
</tr>
</tbody>
</table>

an average weighted according to the number of random sires in the two subsets (for the heritabilities and the genetic correlations) or according to the number of degrees of freedom for the residuals (for residual variances and residual correlations).

Because standard errors could not be computed, approximations were calculated (11). A comparison of these approximate standard errors with empirical results from simulation studies showed poor agreement for parameters accounting for discrete traits. Therefore, only an approximate standard error for the estimated heritabilities of the yield traits is given.

RESULTS AND DISCUSSION

Table 2 gives the frequency of the three disease traits in the data and the phenotypic means for the three production traits in the disease groups. Because of the large amount of data, all differences in yield traits between diseased and undiseased groups were highly significant (P ≤ .001) except those for protein percentage and mastitis (P ≤ .05) and those for fat percentage and any disease (not significant). The mean milk yield of the diseased animals for all three disease traits was considerably higher than the overall mean of the population. This may partly be due to the fact that all lactations shorter than 250 d were discarded, and many cows with diseases were not included because of culling. Diseases are important reasons for culling. In Norway, close to 30% of the culled cows are culled because of disease and low yield.

For dispersion parameters pertaining to one trait only, i.e., the heritabilities \( h_i^2 \) and \( h_j^2 \) and the residual variance of the yield traits \( \sigma_y^2 \), three different estimates were obtained, depending on the pair of traits analyzed at a time. The heritability of the continuous trait milk yield, e.g., was estimated independently from these combinations: milk yield and mastitis, milk yield and ketosis, and milk yield and any disease. For the heritabilities and the residual variances of the yield traits, however, the estimates from the three combinations agreed well; therefore, only one set of estimates is given in Table 3 (heritabilities) and Table 4 (residual variances).

In general, the estimated heritabilities for the yield traits were well within the expected range. The same was true for the estimated residual variances. The standard errors of the estimated heritabilities for the yield traits should be less than .03. For milk yield, the heritability estimates obtained from the two subsets are quite different (.225 versus .308), which may be because of the two samples of test sires used.

The two sets of data were quite different in size and also in genetic background, which might have influenced the results and may

TABLE 2. Frequency of the three disease traits considered and phenotypic means and standard errors of means of the three yield traits for all cows and only the diseased cows.

<table>
<thead>
<tr>
<th>Relative frequency of disease</th>
<th>Milk</th>
<th>Fat</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>(%: X SE)</td>
<td>(kg)</td>
<td>(% SE)</td>
<td>(kg) SE</td>
</tr>
<tr>
<td>All cows</td>
<td>5017</td>
<td>1.02</td>
<td>.001</td>
</tr>
<tr>
<td>Mastitis</td>
<td>5160</td>
<td>3.14</td>
<td>.004</td>
</tr>
<tr>
<td>Ketosis</td>
<td>5209</td>
<td>3.93</td>
<td>.005</td>
</tr>
<tr>
<td>Any disease</td>
<td>5157</td>
<td>2.04</td>
<td>.003</td>
</tr>
</tbody>
</table>
explain some of the differences. Subset 2 was more than twice the size of subset 1. Consequently, subset 2 might reflect a larger part of the total genetic variation. The proportion of imported semen used in both subsets was similar. In subset 1, however, 4 bulls were sires of almost 60% of the test bulls, whereas in subset 2, 7 bulls were sires for a similar fraction of the test bulls.

The estimated heritabilities for the disease traits (Table 5) were very different both with respect to the subset analyzed and with respect to the combination with a yield trait. Differences were up to 115% between subsets 1 and 2 (h_d^2 = .057 versus .123 for any disease when combined with milk yield) and 65% between trait combinations (h_d^2 = .057 versus .094 for any disease in subset 1 when combined with milk yield or protein percentage). In general, the estimated heritabilities of .05 to .12 were in the high region of comparable results from the relevant literature (1).

A simulation study (14) has shown that the heritability of the binary trait is severely overestimated when the average smallest subclass group size is small. This is in agreement with other results (6), in which 2 is given as the critical lower limit for the smallest subclass size. In the present study, the average number of observations per test sire by herd-year-season combination was 1.028. A value of 1.1 in the simulation study led to a 50% overestimation of the heritability value.

Based on these results, the heritability estimates in Table 5 should be interpreted cautiously because they are likely severely overestimated. The variability of the estimates both between subsets and between trait combinations is a further indication of the problems in the methodology, even though there might be different genetic variation for the different disease traits in the two subsets.

Although there is a large impact on the estimated heritability of the binary traits, the estimates of genetic correlations between continuous and binary traits seem to be unaffected by the problems related to the observation structure (14). The standard errors of these estimates empirically were found to be much larger than the expected values for a genetic correlation between two continuous traits estimated from similarly structured data. This may be a reason for the large differences between estimates of r_a (Table 6) from subsets 1 and 2, especially for the milk component traits and both mastitis and any disease, where the estimates were close to 0 from subset 1 and were −2 to −.35 for subset 2. There might be systematic differences between the genetic make-up of the sires in the two subsets as discussed earlier with respect to the differences in the heritability for milk yield. More uniform estimates were obtained for the genetic correlations between milk yield and all three disease traits.

### Table 3. Estimated heritabilities for the yield traits, given separately for the two subsets and pooled according to number of test sires.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Milk yield</th>
<th>Fat percentage</th>
<th>Protein percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.225</td>
<td>.343</td>
<td>.431</td>
</tr>
<tr>
<td>2</td>
<td>.308</td>
<td>.334</td>
<td>.431</td>
</tr>
<tr>
<td>Pooled</td>
<td>.275</td>
<td>.338</td>
<td>.431</td>
</tr>
</tbody>
</table>

### Table 4. Estimated residual variances of the yield traits, given separately for the two subsets and pooled according to number of degrees of freedom for the residual component.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Milk yield</th>
<th>Fat percentage</th>
<th>Protein percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>159,061</td>
<td>.2396</td>
<td>.1095</td>
</tr>
<tr>
<td>2</td>
<td>151,353</td>
<td>.2492</td>
<td>.1074</td>
</tr>
<tr>
<td>Pooled</td>
<td>153,475</td>
<td>.2466</td>
<td>.1080</td>
</tr>
</tbody>
</table>

### Table 5. Estimated heritabilities for the disease traits from different combinations with the three yield traits, given separately for the two subsets and pooled according to number of test sires.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Milk yield</th>
<th>Fat percentage</th>
<th>Protein percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mastitis 1</td>
<td>.053</td>
<td>.055</td>
<td>.055</td>
</tr>
<tr>
<td>2</td>
<td>.111</td>
<td>.089</td>
<td>.070</td>
</tr>
<tr>
<td>Pooled</td>
<td>.088</td>
<td>.076</td>
<td>.064</td>
</tr>
<tr>
<td>Ketosis 1</td>
<td>.075</td>
<td>.098</td>
<td>.107</td>
</tr>
<tr>
<td>2</td>
<td>.079</td>
<td>.111</td>
<td>.106</td>
</tr>
<tr>
<td>Pooled</td>
<td>.078</td>
<td>.106</td>
<td>.106</td>
</tr>
<tr>
<td>Any disease 1</td>
<td>.057</td>
<td>.087</td>
<td>.094</td>
</tr>
<tr>
<td>2</td>
<td>.123</td>
<td>.132</td>
<td>.134</td>
</tr>
<tr>
<td>Pooled</td>
<td>.097</td>
<td>.114</td>
<td>.118</td>
</tr>
</tbody>
</table>
TABLE 6. Estimated genetic correlations between yield traits and disease traits, given separately for the two subsets and pooled according to number of test sires.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Mastitis</th>
<th>Ketosis</th>
<th>Any disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk yield</td>
<td>.472</td>
<td>.655</td>
<td>.574</td>
</tr>
<tr>
<td>2</td>
<td>.330</td>
<td>.652</td>
<td>.563</td>
</tr>
<tr>
<td>Pooled</td>
<td>.507</td>
<td>.653</td>
<td>.567</td>
</tr>
<tr>
<td>Fat percentage</td>
<td>.012</td>
<td>-.483</td>
<td>.065</td>
</tr>
<tr>
<td>2</td>
<td>-.261</td>
<td>-.318</td>
<td>-.210</td>
</tr>
<tr>
<td>Pooled</td>
<td>-.154</td>
<td>-.383</td>
<td>-.102</td>
</tr>
<tr>
<td>Protein percentage</td>
<td>.003</td>
<td>-.671</td>
<td>-.007</td>
</tr>
<tr>
<td>2</td>
<td>-.206</td>
<td>-.630</td>
<td>-.351</td>
</tr>
<tr>
<td>Pooled</td>
<td>-.124</td>
<td>-.646</td>
<td>-.210</td>
</tr>
</tbody>
</table>

traits, which were around -.5. This means that high milk yield is genetically associated with a high susceptibility to the diseases covered by this study. This result may, in addition to what has been mentioned earlier, be a reason for the superiority in milk yield of diseased cows as indicated in Table 2. The estimated genetic correlation between mastitis and milk yield is somewhat higher than what generally is found in literature. In an overview (1), correlations from slightly negative to .66 are cited, most in the area of .20 to .30. The highest one was estimated from a relatively large field data set (8).

In a recent study (2), genetic parameters for mastitis incidence, somatic cell count, and milk yield were estimated from a rather large amount of data using an approximate multiple trait restricted maximum likelihood procedure (10). Because more than 90% of the mastitis observations fell into the category 0 and the remaining observations fell into three more categories (1, 2, and >2), the authors transformed the estimated heritabilities as if observations were of binomial nature (12); estimates on the underlying scale were in the range of .03 to .08. The authors failed to account for the almost binomial observation structure (7) when estimating genetic correlations between mastitis and the continuous variables, somatic cell count and milk yield. The reported genetic correlations between mastitis and milk yield ranged from -.1 to .32 for different regional subsets of the data and, therefore, are expected to have an absolute downward bias.

Distinct negative estimates of $r_s$ were obtained for fat percentage and ketosis (-.38) and protein percentage and ketosis (-.65), which indicates that breeding for high protein and fat percentage should reduce ketosis frequency. Very few estimates of the genetic correlation between ketosis and milk yield can be found in literature (1); most are in the range .20 to .40 and so are also smaller than the ones presented here. The estimated residual correlations given in Table 7 are generally low, the most distinct between milk yield and ketosis (.20) and between protein percentage and ketosis (-.16).

**CONCLUSIONS**

These results are one more piece of information added to the complex puzzle of the relationship of yield and disease in dairy cattle. Like most other studies, the data and methodology used here have certain shortcomings, which we tried to point out clearly. However, it is difficult to assess how the estimates were affected. There is definitely a need for further development of methodology, to allow for missing or incomplete yield data and to enable analyses of multiple traits (e.g., several lactations) simultaneously.

The most striking results of the present study are the high estimated genetic correlations between milk yield and all three disease complexes. If these estimates can be verified in other, independent analyses, consequences for
CORRELATION OF DISEASE AND YIELD TRAITS

dairy breeding programs that ignore information on diseases entirely, have to be considered. The relative economic importance of diseases may be quite different for different populations and disease complexes (9). If we continue placing almost all emphasis on milk yield and closely related traits and do not include disease traits in the breeding goal, consequences may be detrimental for future health and total economic merit of dairy cows.

ACKNOWLEDGMENTS

This study was carried out while H. Simianer was visiting the University of Guelph, supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

17 Tate, R. F. 1955. The theory of correlation between two continuous variables when one is dichotomized. Biometrika 42:225.