

# Mastitis in Dairy Production: Estimation of Sensitivity, Specificity and Disease Prevalence using MCMC

Caroline Uhler

Department of Statistics, University of California, Berkeley

November 23, 2007

## 1 Overview

Mastitis, an endemic disease of dairy cows spread all over the world, is an important cause of loss of efficiency in milk production. The average damage caused by mastitis in Switzerland has been estimated to be CHF 350 per cow and year (Walkenhorst, 2004). Early medical treatment substantially reduces the non-reversible losses in milk production caused by this infection. Therefore, it is essential to use tests with a high probability of detecting mastitis when it is present (i.e. high sensitivity) and of providing a negative result in non-infected cows (i.e. high specificity).

This study investigates the diagnostic efficiency of the four major mastitis diagnostic tests, namely a bacteriological test, the industry standard of somatic cell counting (SCC test), the recently developed test measuring mammary associated amyloid A (MAA test) and a test measuring the electrical conductivity of milk (MEC test). More information about these tests can be found in Hogeveen (2005). For this study 25 cows were tested at all four quarters of the udder with each of the four mastitis diagnostic tests. It would be simple to estimate the sensitivity and specificity of the tests if a perfect test were available for comparison. However, a gold standard for the detection of mastitis does not yet exist. But by introducing the true disease status of the cows as auxiliary variables, the disease prevalence as well as the sensitivities and specificities of the four mastitis diagnostic tests can be estimated by using MCMC.

Under the assumption that the four mastitis diagnostic tests are independent, it is straightforward to get posterior distributions and estimates of the parameters by Gibbs sampling. However, tests that measure similar biological processes are likely to be positively dependent when conditioning on the true disease status. The production of somatic cells and MAA both belong to the body's defense mechanisms, which could lead to a positive dependence between the SCC test and the MAA test. In the following, I will explain the graphical model under these relaxed assumptions. Under this model it is difficult to use just the Gibbs sampler and to sample directly from the full conditionals. Therefore, a Metropolis-Hastings step is introduced within the Gibbs sampler. Finally, model selection is performed by reversible jump MCMC.

## 2 Model

### 2.1 Notation

When applying four tests to one population, the observed data can be classified into a  $2 \times 2 \times 2 \times 2$  contingency table. Each cell holds the count of tested individuals with a given combination of the four binary tests  $T_j$ ,  $j = 1, 2, 3, 4$ , where  $T_1$  denotes the bacteriological test,  $T_2$  the SCC test,  $T_3$  the MAA test and  $T_4$  the MEC test. A negative result on the  $j^{\text{th}}$  test is denoted by  $T_j = 0$  and a positive result by  $T_j = 1$ .

The parameters of primary interest in diagnostic testing are the prevalence of the disease, which is denoted by  $\pi$ , and the specificity and the sensitivity of the different tests, denoted by  $\text{Sp}_j$  and  $\text{Se}_j$ ,  $j = 1, 2, 3, 4$ , respectively.

## 2.2 Dependence of tests

As defined and further discussed in Gardner et al. (2000), two tests are *conditionally independent* when the sensitivity (or specificity) of the second test does not depend on whether the results of the first test are positive or negative among infected (or non-infected) individuals.

As mentioned before, because the SCC test and the MAA test measure similar biological processes, they are likely to be positively dependent when conditioning on the true disease status. The positive conditional dependence between the specificity of the SCC test and the MAA test can be expressed as

$$p_{.00.|\bar{D}} := P(T_2 = 0, T_3 = 0 | \bar{D}) > P(T_2 = 0 | \bar{D}) P(T_3 = 0 | \bar{D}).$$

and the dependence of test sensitivities as

$$p_{.11.|D} := P(T_2 = 1, T_3 = 1 | D) > P(T_2 = 1 | D) P(T_3 = 1 | D)$$

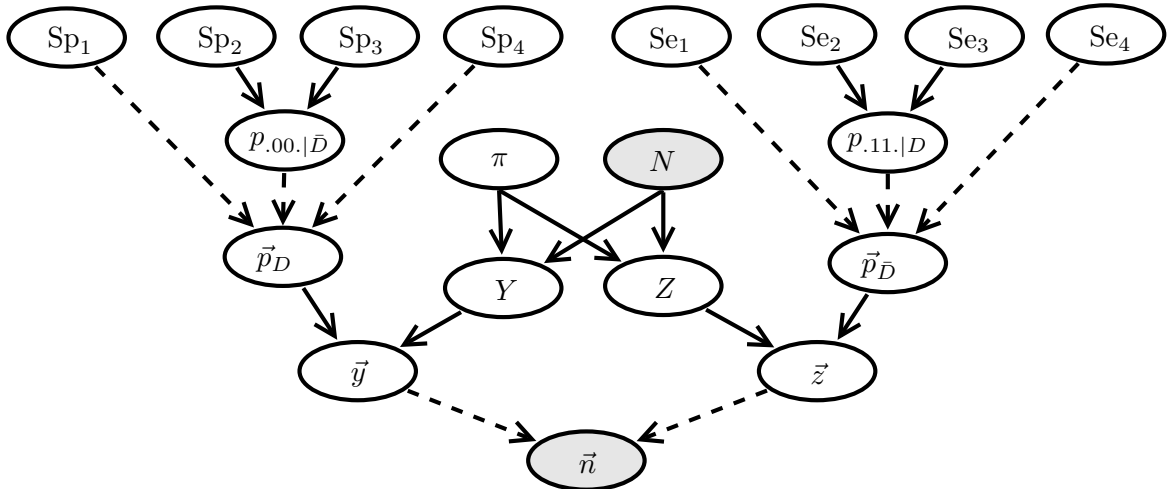
Note that a dependence of test sensitivities does not necessarily imply a dependence of test specificities and vice versa.

Thus, the full conditional dependence model, where test responses are correlated for both disease states, introduces two additional parameters, namely  $p_{.00.|\bar{D}}$  and  $p_{.11.|D}$  with the constraints  $p_{.00.|\bar{D}} > Sp_2 Sp_3$  and  $p_{.11.|D} > Se_2 Se_3$ .

## 2.3 Graphical model

The graphical model under the relaxed assumptions is a directed acyclic graph and is depicted in Figure 1. For the independence model the two additional parameters  $p_{.00.|\bar{D}}$  and  $p_{.11.|D}$  can be omitted. In Figure 1 solid arrows represent probabilistic dependencies, while dashed arrows represent deterministic relationships.

Figure 1: Directed acyclic graph for the independence model.



The model is specified as follows:

$$Sp_t \sim \text{Beta}(a_{Sp_t}, b_{Sp_t}), \quad t = 1, 2, 3, 4$$

$$\pi \sim \text{Beta}(a_\pi, b_\pi)$$

$$Y | N, \pi \sim \text{Binomial}(N, 1 - \pi)$$

$$Z | N, \pi \sim \text{Binomial}(N, \pi)$$

$$p_{.00.|\bar{D}} \sim \text{Uniform}(Sp_1 \cdot Sp_2, \min(Sp_1, Sp_2))$$

$$Se_t \sim \text{Beta}(a_{Se_t}, b_{Se_t}), \quad t = 1, 2, 3, 4$$

$$\vec{n} = \vec{z} + \vec{y}$$

$$\vec{y} | Y, \vec{p}_D \sim \text{Multinomial}(Y, \vec{p}_D)$$

$$\vec{z} | Z, \vec{p}_{\bar{D}} \sim \text{Multinomial}(Z, \vec{p}_{\bar{D}})$$

$$p_{.11.|D} \sim \text{Uniform}(Se_1 \cdot Se_2, \min(Se_1, Se_2))$$

$$\begin{aligned}
p_{ijkl|\bar{D}} &= \text{Sp}_1^{1-i} (1 - \text{Sp}_1)^i (1 - \text{Sp}_2 - \text{Sp}_3 + p_{.00.|\bar{D}})^{j \cdot k} (\text{Sp}_2 - p_{.00.|\bar{D}})^{(1-j) \cdot k} (\text{Sp}_3 - p_{.00.|\bar{D}})^{j \cdot (1-k)} \\
&\quad p_{.00.|\bar{D}}^{(1-j)(1-k)} \text{Sp}_4^{1-l} (1 - \text{Sp}_4)^l \\
p_{ijkl|D} &= \text{Se}_1^i (1 - \text{Se}_1)^{1-i} (\text{Se}_2 - p_{.11.|D})^{j \cdot (1-k)} (\text{Se}_3 - p_{.11.|D})^{(1-j) \cdot k} (1 - \text{Se}_2 - \text{Se}_3 + p_{.11.|D})^{(1-j)(1-k)} \\
&\quad p_{.11.|D}^{j \cdot k} \text{Se}_4^l (1 - \text{Se}_4)^{1-l}
\end{aligned}$$

where  $p_{ijkl|D} = P(T_1 = i, T_2 = j, T_3 = k, T_4 = l | D)$  respectively  $p_{ijkl|\bar{D}} = P(T_1 = i, T_2 = j, T_3 = k, T_4 = l | \bar{D})$ ,  $Z$ , respectively  $Y$ , denote the total number of disease positive individuals, respectively disease negative individuals, and  $N$  the total number of individuals. Analogously,  $\vec{z}$ , respectively  $\vec{y}$ , denote the vectors of counts  $(z_{ijkl})_{i,j,k,l \in \{0,1\}}$ , respectively  $(y_{ijkl})_{i,j,k,l \in \{0,1\}}$ , and  $\vec{n}$  the vector  $(n_{ijkl})_{i,j,k,l \in \{0,1\}}$ .

I chose the beta distribution as prior because its use greatly simplifies calculations. In addition, beta distributions can yield a large array of potential shapes.

In order to compute the full joint distributions, it is important to note that the direct descendants of  $\text{Se}_t$  are not  $z_{ijkl}$ ,  $i, j, k, l = 0, 1$ , as it may seem in Figure 1.  $\text{Se}_t$  has only an influence on test  $t$  and on the truly diseased individuals. Therefore, the children of e.g.  $\text{Se}_1$  are  $z_{1jkl}$ ,  $j, k, l = 0, 1$ , which are distributed  $\text{Binomial}(z_{.jkl}, \text{Se}_1)$ , where the dot subscript indicates the sum over that index. Similarly,  $\text{Sp}_t$  influences only test  $t$  and the truly non-diseased individuals.

With these explanations the full conditional distributions of each model quantity under the independence model can easily be computed:

$$\begin{aligned}
\pi | \Theta_{-\pi}, \vec{n}, \vec{z} &\sim \text{Beta}(a_\pi + Z, b_\pi + N - Z), \\
\text{Se}_1 | \Theta_{-\text{Se}_1}, \vec{n}, \vec{z} &\sim \text{Beta}(a_{\text{Se}_1} + z_{1\dots}, b_{\text{Se}_1} + z_{0\dots}), \\
\text{Sp}_1 | \Theta_{-\text{Sp}_1}, \vec{n}, \vec{z} &\sim \text{Beta}(a_{\text{Sp}_1} + n_{0\dots} - z_{0\dots}, b_{\text{Sp}_1} + n_{1\dots} - z_{1\dots}),
\end{aligned}$$

where,  $\Theta$  denotes the set of all parameters and for any parameter  $\theta \in \Theta$ ,  $\Theta_{-\theta}$  represents the set of all parameters except for  $\theta$ . The full joint distributions of the sensitivities and specificities of the other tests are computed analogously.

For the computation of the full joint distribution of the auxiliary variable  $\vec{z}$ , I need to introduce the probability  $P_{ijkl}$  of being diseased given that an individual has test combination  $\{i, j, k, l\}$ . Bayes theorem yields

$$\begin{aligned}
P_{ijkl} &= P(D | T_1 = i, T_2 = j, T_3 = k, T_4 = l) \\
&= \frac{p_{ijkl|D} \pi}{p_{ijkl|D} \pi + p_{ijkl|\bar{D}} (1 - \pi)}
\end{aligned}$$

and

$$z_{ijkl} | n_{ijkl} \sim \text{Binomial}(n_{ijkl}, P_{ijkl}).$$

However, due to the parameter space restrictions introduced by the additional parameters  $p_{.00.|\bar{D}}$  and  $p_{.11.|D}$  under the relaxed assumptions, it is difficult to use just the Gibbs sampler and to sample directly from the full conditionals. Therefore, a Metropolis-Hastings step is introduced for updating the parameters  $(\text{Se}_2, \text{Se}_3, p_{.11.|D})$  and  $(\text{Sp}_2, \text{Sp}_3, p_{.00.|\bar{D}})$ .

If there were no restrictions on the probability vector  $\vec{p}_D := (p_{.00.|D}, p_{.10.|D}, p_{.01.|D}, p_{.11.|D})$ , respectively on the vector  $\vec{p}_{\bar{D}} := (p_{.00.|\bar{D}}, p_{.10.|\bar{D}}, p_{.01.|\bar{D}}, p_{.11.|\bar{D}})$ , and a Dirichlet(1, 1, 1, 1) prior was used, the full conditional distribution would be Dirichlet( $z_{.00.} + 1, z_{.10.} + 1, z_{.01.} + 1, z_{.11.} + 1$ ), respectively Dirichlet( $n_{.00.} - z_{.00.} + 1, n_{.10.} - z_{.10.} + 1, n_{.01.} - z_{.01.} + 1, n_{.11.} - z_{.11.} + 1$ ). These distributions are used as proposal distributions in the Metropolis-Hastings algorithm. So a proposed probability vector  $\vec{p}_D^*$  is generated from the full conditional Dirichlet distribution and the proposed sensitivities,  $\text{Se}_2^* = p_{.11.|D}^* + p_{.10.|D}^*$  and  $\text{Se}_3^* = p_{.11.|D}^* + p_{.01.|D}^*$  are computed. If  $p_{.11.|D}^* > \text{Se}_2^* \text{Se}_3^*$  the chain moves to the new proposed parameters with probability

$$P(\text{move}) = \min \left( 1, \frac{\min(\text{Se}_2, \text{Se}_3) - \text{Se}_2 \text{Se}_3}{\min(\text{Se}_2^*, \text{Se}_3^*) - \text{Se}_2^* \text{Se}_3^*} \prod_{t=2}^3 \left( \frac{\text{Se}_t^*}{\text{Se}_t} \right)^{a_{\text{Se}_t} - 1} \left( \frac{1 - \text{Se}_t^*}{1 - \text{Se}_t} \right)^{b_{\text{Se}_t} - 1} \right).$$

Otherwise, the parameters are not updated in this iteration. Updating  $p_{.00.|\bar{D}}$ ,  $Sp_2$  and  $Sp_3$  works analogously.

Herewith, the independence and the full conditional dependence model are constructed. Updating the partial dependence models is accomplished analogously. I account for the uncertainty about the dependence structure of the SCC test and the MAA test by running reversible jump MCMC on the four possible dependence models.

## 2.4 Reversible jump MCMC

Reversible jump MCMC is performed on the following four models, where every model is a priori assumed to be equally likely.

$$\begin{aligned} M_1 : \quad & p_{.11.|\bar{D}} = Se_2 Se_3, & p_{.00.|\bar{D}} &= Sp_2 Sp_3 \\ M_2 : \quad & p_{.11.|\bar{D}} > Se_2 Se_3, & p_{.00.|\bar{D}} &= Sp_2 Sp_3 \\ M_3 : \quad & p_{.11.|\bar{D}} = Se_2 Se_3, & p_{.00.|\bar{D}} &> Sp_2 Sp_3 \\ M_4 : \quad & p_{.11.|\bar{D}} > Se_2 Se_3, & p_{.00.|\bar{D}} &> Sp_2 Sp_3 \end{aligned}$$

- First, consider a move from  $M_1$  to  $M_2$  or from  $M_3$  to  $M_4$ . For this,  $\gamma$  is sampled from the uniform distribution  $U(0, \min(Se_2, Se_3) - Se_2 Se_3)$  and  $p_{.11.|\bar{D}}^* = Se_2 Se_3 + \gamma$  is proposed. The acceptance probability for a move from  $M_1$  to  $M_2$  or from  $M_3$  to  $M_4$  is

$$\begin{aligned} P(\text{move}) &= \min \left( 1, \frac{p(\vec{n} | M_2, \Theta_2) p(\Theta_2 | M_2) p(M_2)}{p(\vec{n} | M_1, \Theta_1) p(\Theta_1 | M_1) p(M_1)} \frac{1}{(\min(Se_1, Se_2) - Se_1 Se_2)^{-1}} \right) \\ &= \min \left( 1, \frac{P(\vec{z} | Se_1, Se_2, p_{.11.|\bar{D}})}{P(\vec{z} | Se_1, Se_2)} \right), \end{aligned} \quad (1)$$

where  $\Theta_1 = \{Se_2, Se_3\}$ ,  $\Theta_2 = \{Se_2, Se_3, p_{.11.|\bar{D}}\}$ . For the second equality I have used that the prior probability of each model is  $\frac{1}{4}$ , that  $\Theta_1$  and  $\Theta_2$  only affect the truly diseased individuals and that

$$p(\theta_2 | M_2) = p(p_{.11.|\bar{D}} | Se_2, Se_3, M_2) p(\theta_1 | M_1).$$

- For the reverse moves from  $M_2$  to  $M_1$  or from  $M_4$  to  $M_3$  define  $p_{.11.|\bar{D}}^* = Se_1 Se_2$ . The acceptance probability is computed analogously to (1):

$$P(\text{move}) = \min \left( 1, \frac{P(\vec{z} | Se_2, Se_3)}{P(\vec{z} | Se_2, Se_3, p_{.11.|\bar{D}})} \right).$$

- For moving between  $M_1$  and  $M_3$  or between  $M_2$  and  $M_4$  similar proposals are used.

So one iteration of the algorithm consists of first updating the parameters and then jumping between the models. The whole process is iterated until convergence.

## 3 Data Analysis

For this study 25 cows from a Swiss farm were tested at all four quarters of the udder with the four mastitis diagnostic tests. As the four quarters of a cow are individual units, it can be assumed that 25 cows result in 100 independent samples. The test results are summarized in the following  $2 \times 2 \times 2 \times 2$  data matrix:

$$\begin{array}{cccc} n_{0000} = 41 & n_{1000} = 7 & n_{0100} = 1 & n_{1100} = 1 \\ n_{0010} = 2 & n_{1010} = 2 & n_{0110} = 2 & n_{1110} = 3 \\ n_{0001} = 26 & n_{1001} = 8 & n_{0101} = 0 & n_{1101} = 0 \\ n_{0011} = 0 & n_{1011} = 0 & n_{0111} = 1 & n_{1111} = 6. \end{array}$$

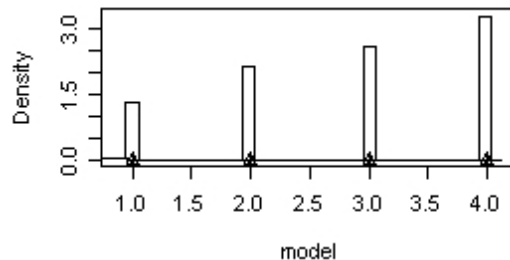
Because no prior information about the model quantities was available, I used  $\text{beta}(0.5, 0.5)$  priors for the sensitivities and specificities of the four tests as well as for the prevalence. I generated three Markov chains with a length of 9000 iterations with different starting values and used the tools described in Toft et al. (2007) and in Gilks et al. (1995) to assess convergence of the chains. This included plotting the running mean of multiple sequences with overdispersed starting points, analyzing the Gelman-Rubin statistic and the autocorrelations. After discarding an initial burn-in of 2000 iterations, only every seventh sample of each chain was saved due to high autocorrelations. These samples of the three chains were then combined for the summary statistics.

The summary statistics of the combined chains are shown in Table 1 and the posterior distribution of the parameter ‘model’ in Figure 2.

Table 1: Summary statistics of the combined chains. The sample means and standard deviations are given in the first two columns. Quantiles are given in the following columns.

	mean	sd	2.5%	25%	50%	75%	97.5%
$\pi$	0.17	0.07	0.07	0.13	0.16	0.20	0.33
$\text{Se}_{\text{Bact}}$	0.75	0.13	0.49	0.67	0.76	0.86	0.98
$\text{Se}_{\text{SCC}}$	0.76	0.17	0.34	0.59	0.73	0.84	0.97
$\text{Se}_{\text{MAA}}$	0.77	0.16	0.39	0.67	0.80	0.90	0.99
$\text{Se}_{\text{MEC}}$	0.51	0.14	0.25	0.41	0.51	0.60	0.79
$\text{Sp}_{\text{Bact}}$	0.82	0.05	0.71	0.79	0.82	0.85	0.94
$\text{Sp}_{\text{SCC}}$	0.96	0.03	0.89	0.95	0.97	0.98	1.00
$\text{Sp}_{\text{MAA}}$	0.94	0.03	0.87	0.93	0.95	0.97	1.00
$\text{Sp}_{\text{MEC}}$	0.61	0.05	0.59	0.57	0.61	0.64	0.71
$p_{.11. D}$	0.61	0.17	0.24	0.49	0.63	0.74	0.91
$p_{.22. \bar{D}}$	0.92	0.04	0.83	0.90	0.93	0.95	0.98

Figure 2: Posterior density of the parameter ‘model’.



## 4 Discussion

The farm analyzed in this study shows a quarter prevalence of 17%. In other studies, the quarter prevalence of mastitis in Europe was estimated to be about 20%. So the farm analyzed in this study lies near the average.

The MEC test is cheap to perform (USD 0.10 per test), but the sensitivity and the specificity of the MEC test are very low. The high probability of false negatives and false positives is due to the fact that the electrical conductivity is influenced by mastitis infection but also by many other factors such as fat content and temperature of the milk, milking intervals and food types. The bacteriological test achieves

only moderately good results and is highly expensive to perform (USD 10 per test). The MAA test and the SCC test have the same moderate costs (USD 1 per test) and both show a very good test accuracy.

The summary statistics reveal a relatively high standard deviation of the sensitivities and show that the estimates of the sensitivities are less accurate than the estimates of the specificities. This can be explained by the fact that the sensitivities are computed from the truly infected quarters, the specificities from the non-infected quarters. The quarter prevalence is 17% and therefore only about 17 quarters are truly infected. This is too small a sample size to achieve good estimates with a small standard deviation. In the future, this analysis should be repeated with a larger data set.

The posterior density of the parameter 'model' shows that the chains tend to remain longer in the full conditional dependence model. This suggests that the SCC test and the MAA test are conditionally dependent. However, further analysis has shown that this tendency is not significant. So I can conclude that all tests are independent.

The sensitivities of the four tests are lower than the specificities. This suggests that the tests could be improved by choosing a lower discrimination threshold for a positive test result.

Further research should include the computation of the ROC curves in order to compare the overall accuracy of the mastitis diagnostic tests independently of the chosen decision threshold. For the future, I also strongly recommend to collect data upon the costs of a true/false positive/negative test result to estimate the economic effects and find the overall optimal decision thresholds in order to minimize the economic damage of mastitis in dairy production.

## References

- Gardner, I. A., Stryhn, H., Lind, P., Collins, M. T. (2000). Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Preventive Veterinary Medicine*, 45, 107–122.
- Gilks, W. R., Richardson, S., Spiegelhalter, D. J. (Eds) (1995). *Markov chain Monte Carlo in practice*. Chapman & Hall, London.
- Hogeveen, H. (Ed.) (2005). *Mastitis in dairy production*. Wageningen Academic Publishers, The Netherlands.
- Toft, N., Innocent, G. T., Gettinby, G., Reid, S. W.J. (2007). Assessing the convergence of Markov Chain Monte Carlo methods: An example from evaluation of diagnostic tests in absence of a gold standard. *Preventive Veterinary Medicine*, 79, 244–256.
- Walkenhorst, M. (2004). Eine gute Prävention erhält die Eutergesundheit. *Die Grüne*, 4, 38–39.