An individual modelling tool for within and between lactation consecutive cases of clinical mastitis in the dairy cow: an approach based on a survival model

Patrick GASQUIa*, Jean-Baptiste COULONb, Odile PONSc

a Unité d'Épidémiologie Animale, INRA, 63122 Saint-Genès-Champanelle, France
b Unité de Recherche sur les Herbivores, INRA, 63122 Saint-Genès-Champanelle, France
c Unité de Biométrie, INRA, 78352 Jouy-en-Josas, France

(Received 21 December 2001; accepted 22 August 2002)

Abstract – Clinical mastitis in dairy cows has for many years been the subject of numerous epidemiological surveys to determine the main risk factors. In most cases this data has been analysed using a standard Poisson model without taking into consideration possible dependence between consecutive pathological events. These analyses have brought to light a great many potential risk factors without making it possible to clarify a certain amount of confusion surrounding the effects. The extension of an individual within a lactation model, considering dependence between clinical cases of mastitis within lactation so as to take into account inter-lactation dependence (which has already been published) is presented in the form of mixed distributions within the same survival model framework. By introducing new parameters, infection rate at calving and the identification of a higher exogenous infection rate indoors than at pasture, it is possible to take into consideration what had previously appeared to be a lactation stage factor, a calving month factor or even part of a parity factor. By considering these two types of dependence within the same model, it appears to be possible to obtain a simpler model in terms of the factors to be taken into account, and one that is based on generally acknowledged and easily understandable biological considerations. Lastly, a possible way of extending the model is to consider the dry period before calving and this is presented. It would make it possible to envisage developing a complete model of the animal's lifetime in the not-too-distant future. It is still necessary, however, to determine the farming system factors in the general sense of the term, which specifically affect one or the other of the different model parameters, before one can draw conclusions as to the potential extension of this type of model. A national survey is currently being carried out on approximately 600 French breeding farms that will help meet this last objective.

clinical mastitis / individual model / dairy cow / survival model / recurrence

* Correspondence and reprints
Tel.: (33) 4 73 62 42 65; fax: (33) 4 73 62 45 48; e-mail: Patrick.Gasqui@clermont.inra.fr
1. INTRODUCTION

In the field of animal epidemiology, clinical mastitis in dairy cows has for many years been the subject of a great number of surveys aimed at determining the main risk factors [4–7, 9, 12, 24, 25, 27, 29, 33–35, 37, 41]. In most of these studies, the elements used for diagnosing clinical mastitis have been relatively similar and based mainly on observations of the first jets of milk during milking with identified lumps, associated with the state of the udder [6, 9, 12, 22, 24, 25, 37, 40]. These risk factors were for the most part demonstrated by adjusting the number of cows observed as having clinical mastitis to a Poisson distribution within the framework of a Generalised Linear Model (GLM) type model [28] including fixed effects or even individual random effects. During these analyses, the authors retained only the clinical cases of mastitis arising a certain number of days after a previous episode of clinical mastitis; this time-lapse varies from 8 [22], 10 [24], 14 [1, 4, 12], 30 [13, 40] to 90 [25] days depending on the author. This selection carried out between the mastitis cases observed is aimed at eluding a possible dependence between consecutive events. Some authors even went so far as to consider only the first occurrence of clinical mastitis during lactation in an animal [9, 22, 41]. Similarly, during these analyses, some authors only used one lactation per animal so as to avoid possible dependence between consecutive lactations in a same animal [6, 22, 35, 41]. It is obvious that these different selections can influence whether the statistical significance of such and such a factor included a priori in the model is revealed or is absent.

In a previous article we saw that it would be easy to take the first level of dependence between consecutive clinical cases of mastitis within lactation into consideration in the framework of a mixed survival distribution [17]. This approach gave much better results than that based on a Poisson distribution in the framework of a GLM model, with or without individual random effects [16, 17]. This approach also made it possible to demonstrate a certain number of relatively standard risk factors [17]. It is obvious that, in general, if the model is not statistically correct, the significance tests of the different factors studied are no longer correct; this is without mentioning the many instances of confusion surrounding factors typically encountered in this type of epidemiological analysis. In the previous study [17], although the model appears correct by providing good prediction and validation results, and although it makes it possible to bring to light the factors usually identified, it remains nonetheless incomplete in that it does not take into consideration possible dependence between consecutive lactations in a same animal.

The aim of this article was to set out how to extend this individual model so as to take this second level of dependence into consideration: the possible dependence between consecutive lactations in a same animal. This new model makes it possible to integrate, at the same time, an important factor that had not yet been taken into consideration in the previous model: the animal's circumstances whether indoors or at pasture.

With a view to making estimations and predictions, new concepts of methodology were introduced that were selected to take into consideration potential dependence between consecutive lactations in a same animal vis-à-vis clinical mastitis. These will complement the methodological elements introduced in the first article concerning intra-lactation dependence [17]: clinical mastitis due to exogenous infection (REX) or persistent endogenous infection (REN) [1, 11, 15, 17, 32, 42]. We will then present the results obtained on three experimental farms.
2. MATERIALS AND METHODS

2.1. Data

The data used to validate the model of relationships between consecutive events within lactation and between consecutive lactations from three experimental sites run by INRA (National Institute for Agricultural Research) were the following: (1) Orcival in the Auvergne region with data from 1979 to 1989 that made the development of the model with respect to estimating the parameters possible, (2) Marce- nat in the Auvergne region with data from the 1997–1999 campaign, and Rennes in the Brittany region with data from 1985 to 1989 that made the validation of the model with respect to prediction possible.

On the first site, the farming conditions were relatively homogeneous over the study period: a milking parlour was used, there were 5 months spent indoors under loose housing (from 1 November to 31 March) and the rest of the time was spent at pasture. Three breeds were mainly represented at this site: the Holstein breed (HO), the Holstein-Française Frisonne crossbreed (HO*FF) and the Mont-béliarde-Pie Noire crossbreed (MO*PN).

The farming conditions in the second site were relatively similar to those of the first: a milking parlour was used, there were also 5 months spent indoors under stall-type stabling (from 1 November to 31 March) and the rest of the time was spent at pasture. Three breeds were represented at this site: the Holstein breed (HO), the Mont-béliarde breed (MO) and the Tarine breed (TA); this made it possible to validate the model on the basis of more recent data. At the third site, a milking parlour was used, there were 6.5 months spent indoors (from 1 October to 15 April) and the rest of the time was spent at pasture. Two breeds were represented at this site: the Holstein breed (HO) representing 96% of the lactations, and the crossbreed Holstein-Française Frisonne (HO*FF); this made it possible to validate the model on the basis of data from another region.

At the Orcival site and for the multiparous cows, the time interval between the calving date and the date the cow went back indoors was below 30 days for 28% of the lactations, between 30 and 60 days for 17% of the lactations and above 60 days for 43% of the lactations (respectively 36%, 27% and 34% for the Marce- nat site, 31%, 27% and 39% for the Rennes site). At the Orcival site and for primiparous cows, the results obtained were respectively 64%, 12% and 16% (respectively 48%, 28% and 20% for Marcenat, and 43%, 37% and 14% for Rennes).

2.1.1. Selection of animals and lactations

The aim of the study was to model both the relationship between consecutive cases of mastitis between consecutive lactations and within lactation. Therefore, at the first site, we selected animals for which we had access to data for a set of consecutive lactations after the first one. For the other two sites, we kept all the animals with all the available lactations. For all three sites, no minimal or maximum production time criteria were used. All the lactations were used whatever the calving month. The existence of a disease at calving (other than mastitis) affecting the udder area and known to be potentially associated to the occurrence of clinical mastitis (mammary oedema, hypocalcaemia [10, 21] or even others [18] that occurred less frequently or were absent from the lactations considered) was not used as a sampling criterion. This information was not included in the model since it had not been identified as a significant factor in the previous model [17].

At the second site, the fact that a maximum of only two calving campaigns was available per animal meant that the potential influence of the possible level of dependence between animals of a same herd was reduced considerably. The
hypothesis that independence exists between different animals of a same herd still remains valid a priori for all three sites, since not only did the clinical cases of mastitis observed during the two periods exhibit no epidemic characteristics, but also there was an absence of “peaks” of mastitis cases grouped together over a period of a few days in the three herds studied.

2.1.2. Natural clinical cases of mastitis

In this study, we had access to all the natural clinical cases of mastitis that had been observed and diagnosed on the basis of clinical symptoms (appearance of the milk with at least the presence of lumps in the first jets of milk at milking, status of the quarters and of the udder). As long as the teat after treatment produced clinical symptoms, it was not a new case. Yet, if clinical symptoms were observed in another teat of the same udder, it was a new case for the udder. All the clinical cases of mastitis were kept regardless of the time interval separating the two consecutive occurrences at the udder level, since the model also takes into consideration the possible relationship between consecutive cases of mastitis within lactation.

Therefore, for the model construction data for the Orcival site, we had access to a sample of 693 lactations concerning 247 cows having been affected by 354 clinical cases of mastitis (481 lactations with 0 clinical cases of mastitis, 130 with 1, 45 with 2, 22 with 3, 11 with 4, 2 with 5, 1 with 6 and 1 with 8). For the validation data, we also had, for the Marcenat site, 226 lactations with a maximum of two lactations per animal, concerning 146 cows having been affected by 103 clinical cases of mastitis (168 lactations with 0 clinical cases of mastitis, 35 with 1, 13 with 2, 2 with 3, 6 with 4, 1 with 5 and 1 with 7) and, for the Rennes site, 383 lactations concerning 207 cows having been affected by 249 clinical cases of mastitis (240 lactations with 0 clinical cases of mastitis, 88 with 1, 30 with 2, 15 with 3, 3 with 4, 3 with 5, 3 with 6 and 1 with 11).

2.1.3. Choice of factors

In order to construct the model on the basis of data from the first site, we introduced into the model a certain number of potential factors that were mainly individual and used the factors demonstrated using the intra-lactation model that is described in a previous article [17]: a breed factor with the 3 breeds available, a calving month factor divided into 7 modalities (October, November, December, January, February, March and April to September) and a parity factor divided into 5 modalities (parity 1, 2, 3, 4 and 5 and more). At the Orcival site, 36% of the lactations corresponded to primiparous cows (respectively 29% for the Marcenat site and 54% for the Rennes site). Since it was of interest to study both primiparous and multiparous cows, it was not possible to introduce any factor directly related to the production level into the model, or use such a factor as a selection criterion. Its possible effect would be confused with the parity factor and the breed factor. For the Orcival site, the initial production of 91% of primiparous cows was below 22 kg, whereas the initial production of 77% of multiparous cows was above 22 kg (88% and 90% respectively for the Rennes site). The productive period was therefore divided into 8 consecutive periods: “from calving until day 3”, “from day 4 until day 30”, “from day 31 until day 90”, “from day 91 until day 150”, “from day 151 until day 210”, “from day 211 until day 300”, “from day 301 until day 360” and “beyond the 361st day until the maximal drying off date”. The first seven periods were
respectively 3, 27, 60, 60, 60, 90 and 60 days long.

Two new factors were introduced into the within-between lactation model: an individual factor depending on the time specifying the animal's circumstances whether indoors or at pasture, and an individual factor specifying whether the previous lactation in a multiparous cow was free from clinical mastitis or not. For the first factor, the entry date in the cowshed for the winter and the exit date at pasture in the spring were considered for each cow, depending on the animal calving date for a cow and climate year considerations for a herd. It was a time depending factor for an animal.

2.2. Modelling the relationship between consecutive lactations

In a previous article [17], the importance and legitimacy of the Poisson distribution for approaching the observed distribution of the number of mastitis episodes per lactation [20, 36] was discussed; the intervals between consecutive cases of mastitis were thereby distributed exponentially. The survival model approach [31] makes it possible in particular to take into consideration the productive periods effectively observed and the individual characteristics and farming system characteristics responsible for the different risk levels. It was thereby demonstrated that, during lactation, the risk of "REX" can be considered to be not constant due, for example, to physiological factors (greater risk at the very beginning of lactation and then variable risk thereafter depending on the length of time since calving); we suggest here that environmental factors be taken into account such as the animal's "circumstances" whether indoors or at pasture with a greater risk a priori indoors than at pasture. By taking the risk to be simply constant during consecutive periods within lactation, we obtain a homogeneous Poisson process for fixed intervals. The values of these covariates can remain constant from one lactation to another (post-calving periods and following production) or can be variable (for example depending on the time spent indoors or at pasture).

In this previous study [17], the main element established was the advantage of carrying out modelling based on a mixed exponential distribution in the framework of survival models in order to take into consideration the relationship between consecutive cases of mastitis within lactation. To take into consideration the relationship between consecutive lactations, the approach suggested consists of using the previous model for the relationship between consecutive cases of mastitis within lactation as a base, introducing a new parameter characterising the health status of the udder at calving, at the beginning of lactation; this is so as to take into account what some authors have observed [3, 8, 19, 39]. In a certain proportion \(1 - q\) of cases, the animal calves with a "non-infected" udder. In a proportion \(q\) of cases, the animal calves with an udder that is already in a state of infection; the udder will have become infected either during a previous productive period or during a previous "dry" period. In the first case, the occurrence of at least one episode of clinical mastitis during the previous productive period will potentially be one of the important components of parameter "\(q\)"; some authors have thereby demonstrated the positive correlation between the occurrence of clinical mastitis during the productive period and the occurrences of clinical mastitis having occurred during the previous productive period [9, 35]. It is understood that if an indicator characteristic of the "chronic infection" state of an animal was available, as could be the case for an indicator such as "cell count level at a given moment" in the previous lactation for example, this indicator would naturally take into consideration the information provided by the possible occurrence of a clinical case of mastitis in the previous lactation. In the second case, the occurrence of infections that are not cured...
spontaneously during the dry period before calving will also tend to increase parameter “q”; some authors thereby consider numerous clinical cases of mastitis having occurred during production to be due to infections that were not cured spontaneously having been contracted previously during the dry period before calving [8, 39]. Some authors even go so far as to believe that some of these infections can persist throughout lactation and cannot be cured with the systematic application of a treatment during the drying-off period [30]; parameter “q” could therefore also depend on the level of infection during previous dry periods. By incorporating this new parameter “q”, the model presented diagrammatically in Figure 1 is complete, and makes it possible to simultaneously take into consideration the possible relationships between consecutive episodes of clinical mastitis in a same animal, within lactation and from one lactation to the next throughout its lifetime. Straightaway, one can expect this model to obtain a parameter “q” that will depend on parity number, if only because the first calving is not preceded by drying-off and therefore neither by the previous application of any systematic treatment. Similarly, since the udder’s “chronic infection” status is not easy to “cure”, as is the case in the situation of chronic infection of the udder by Staphylococcus aureus-type germs, one can expect a priori that, in a same animal, the parameter “q” increases with parity number.

The modification made to the previous model, that consisted of adding a mixed term between the calving date and the occurrence of the first case of mastitis during the lactation or the following drying-off or culling date, does not fundamentally modify the methodology used to calculate estimators. This is all the more true since the notions of risk of REX remain unchanged. Only those co-factors or variables with a potential effect a priori on parameter “q” are integrated into the model in order to test how real their effect is.

**Figure 1.** Diagram showing the modelling elements making it possible to take into account a relationship between consecutive cases of mastitis within lactation and between consecutive lactations in the same animal.
For a given lactation, if \( n > 0 \) clinical cases of mastitis are observed at times \( t_1 < t_2 < \ldots < t_{n-1} < t_n \) such that the calving date \( = t_0 < t_1 \) and \( t_n < t^* \) = the animal’s drying-off or culling date, and if the lactation is divided up into \( k \) periods for which the risk of “REX” is considered constant (lactation stage and the animal’s circumstances whether indoors or at pasture), the likelihood then corresponds to this new model in the form:

\[
L(\lambda_1, \ldots, \lambda_k, \theta, \alpha, \gamma)
= \frac{(1-q) \cdot f(t_1-t_0) + q \cdot f_{X_1}(t_1-t_0)}{\prod_{i=2}^{n} \left[ (1-p) \cdot f(t_i-t_{i-1}) + p f_{X_1}(t_i-t_{i-1}) \right]} \cdot \left( \frac{(1-p) \cdot S(t^*-t_0) + p \cdot S_{X_1}(t^*-t_0)}{(1-q) \cdot S(t^*-t_0) + \lambda_{X_1} (t^*-t_0)} \right).
\]

When \( n = 0 \), i.e. when there are no clinical cases of mastitis during lactation, we get:

\[
L(\lambda_1, \ldots, \lambda_k, \theta, \alpha, \gamma)
= \frac{(1-q) \cdot f(t_1-t_0) + q \cdot f_{X_1}(t_1-t_0)}{\prod_{i=2}^{n} \left[ (1-p) \cdot f(t_i-t_{i-1}) + p f_{X_1}(t_i-t_{i-1}) \right]} \cdot \left( \frac{(1-p) \cdot S(t^*-t_0) + p \cdot S_{X_1}(t^*-t_0)}{(1-q) \cdot S(t^*-t_0) + \lambda_{X_1} (t^*-t_0)} \right).
\]

If “\( d \)” is the time interval between two consecutive events (calving, mastitis or drying-off), in the two formulations, the two functions \( f(d) \) and \( S(d) \) are respectively the density function and the survival function of an exponential distribution that remains constant for fixed intervals, whereas the two functions \( f_{X_1}(d) \) and \( S_{X_1}(d) \) are respectively the function of density and the function of survival of an exponential distribution of constant parameter \( \lambda_{X_1} \). In the expression of \( L \), the function \( f \) and \( S \) depend on the duration time between successive mastitis as well as the time of all the previous ones through a step-wise function \( \lambda \).

For a set of \( n^\ast \) lactations, the total likelihood is expressed simply as the product of likelihoods of each of the lactations, i.e.:

\[
L(\theta, \alpha, \beta, \gamma) = \prod_{i=1}^{n^\ast} L_i.
\]

Vector \( \theta \) corresponds to the model’s coefficients that translate the effects of the different factors including the period factor and the environmental factor (indoors/pasture), influencing the value of the “REX” risk with the relationship: \( \lambda = \exp(X_1 \cdot 0) \), where \( X_1 \) is the associated incidence matrix. The exponential relationship ensures that the risk thereby calculated is positive. Vector \( \beta \) corresponds to the model’s coefficients that translate the effects of the different factors influencing the value of the REN rate with a similar relationship: \( \lambda_r = \exp(X_2 \cdot \beta) \), where \( X_2 \) is the associated incidence matrix. Vector \( \alpha \) corresponds to the model's coefficients that translate the effects of the different factors influencing the value of the REN rate with the relationship: \( p = \exp(X_3 \cdot \alpha) / \left[ 1 + \exp(X_3 \cdot \alpha) \right] \), where \( X_3 \) is the associated incidence matrix. This last relationship ensures that the value of parameter \( p \) is therefore between 0 and 1, which corresponds well to the variation range of a probability. Vector \( \gamma \) corresponds to the model’s coefficients that translate the effects of the different factors influencing the value of the infection rate at calving with the relationship: \( q = \exp(X_4 \cdot \gamma) / \left[ 1 + \exp(X_4 \cdot \gamma) \right] \), where \( X_4 \) is the associated incidence matrix. This last relationship ensures that the \( q \) value obtained is between 0 and 1, which corresponds well to the variation range of a probability. This total likelihood depends therefore only on the unknown coefficients \( \theta \), \( \beta \), \( \alpha \) and \( \gamma \) thereby defined. By maximising the total likelihoods according to the model’s unknown coefficients, it is possible to obtain estimates for these coefficients.

If the following model parameters are known: risk of REX, risk of REN, REN rate and infection rate at calving, it is possible to express the law of probability of cases of mastitis observed over any lactation period starting from the calving date in the form of a recurrence; the formulae are similar to those presented in the article cited above [17]. One of the advantages of the approach proposed is that it also makes it possible to use parameter estimations to calculate the distribution function of the occurrence
times of the $w$th case of mastitis within lactation as of the calving date, i.e. $P(T_w \leq t)$ for every time $t > t_0$. In fact $P(T_w \leq t) = P(N(t, t) \geq w) = 1 - P(N(t, t) < w)$ can be deduced from the recurrence formulae modified in the same way as for the likelihood compared to those already presented in the previous article [17].

2.3. Estimation of the parameters modelling the relationship between consecutive cases of mastitis

In order to implement the within-between lactation model, as was the case for the previous intra lactation model [17], adjustments were made using a function from the Fortran mathematical library NAG “Numerical Algorithms Group” (NAG Ltd, Oxford, UK, Mark 16 for Unix) which uses an explicit likelihood function and the first and second derivative values of the coefficients obtained using finite differences. An estimation of the coefficient variance-covariance matrix is obtained from estimates of the inverse of the Fisher information matrix; the Hessian matrix is estimated using a NAG function using finite differences. The need to obtain a REN risk that is strictly higher than the REX risk for any period $k$ of lactation ($k = 1, ..., K$) can, in certain cases, require the use of a maximisation function under the constraint of unequal parameters.

Since the estimators of the coefficients and the linear combinations of these coefficients were constructed by maximum likelihood, their distribution converges towards Gaussian laws under hypotheses that the model is identifiable [26]. These limits mean that it is possible to construct confidence intervals associated with the estimated coefficients or with their linear combinations. Confidence intervals for the model parameters (REX risks, REN rate and infection rate at calving) are deduced by transformation on the basis of the estimations of the associated coefficients; these confidence intervals are generally non-symmetrical.

Sub-model tests can be carried out using the statistic of the likelihood ratio

$$\Lambda = -2 \cdot \log(L_0/L_1)$$

which follows an $\chi^2$ with a degree of freedom ($q_0 - q_1$) if $L_0$ is the likelihood of the general model with $q_0$ coefficients, and if $L_1$ is the likelihood of the sub-model with $q_1$ coefficients. These statistics make it possible to test the effects of different factors of interest while at the same time being consistent with the estimator of the maximum likelihood used to estimate the coefficients and therefore the parameters of each of the models. The parameter estimations are obtained with the Orcival data set.

2.4. Validation of the model including the relationship between consecutive cases of mastitis

The complete model is validated using Martingale residuals, residuals that are usually used for survival models [2, 14] and the formulation of which has been presented explicitly in the publication cited above [17]. In this approach, the observation corresponds to a length of time that ends either with a case of mastitis or with the end of a lactation period without mastitis (right censored) for each lactation observed. When the observation is censored, the residual is strictly negative. By examining the Martingale residuals according to the modalities of the different factors of interest or the order number of the observations organised by time, it is possible to validate the model if the mean of these residuals is zero (absence of systematic bias) and if there are no isolated influential points.

For any lactation, and under the hypothesis that animals within a same herd are independent, the estimated distribution of the number of cases of mastitis per lactation is expressed in exactly the same way as described previously [17] and should be compared to the distribution observed.

The model is also validated with Marcenat and Rennes data sets, in using Orcival parameter estimations.
3. RESULTS

3.1. Estimation of the model parameters using animals from the Orcival site

The complete model – which had been introduced previously – takes into consideration the relationships between consecutive clinical cases of mastitis within lactation and from one lactation to another, the dividing-up of lactations into 8 consecutive periods following calving, the environmental factor which is time-dependent (animal indoors or at pasture), and the more standard factors (parity, calving month and breed). With this complete model, the Martingale residuals are zero on average whatever the modality of the potential factors, and vary between $-1.0$ and $+1.0$, which makes it possible to validate this complete model.

On the basis of the complete model and using consecutive sub-model tests (“forward selection”), we used a selected confidence coefficient a priori of 0.95 to eliminate the factors “breed”, “calving month” and “lactation stage” and to retain only three factors: the parity factor divided into two modalities “primiparous versus multiparous” and the “clinical mastitis during the previous lactation” factor, with these two factors only affecting the infection rate at calving. A third factor “animal indoors or at pasture” was also kept, since it only affected the risk of REX. The final model thus obtained had 7 degrees of freedom and its “P-value” compared to the complete model was 0.92. In this model, the risk of REX only depended on the environmental factor with two modalities (indoors and pasture), and the infection rate at calving depended on the parity factor divided into two modalities (primiparous versus multiparous) – and for the multiparous cows it depended on the presence or absence of at least one clinical case of mastitis during the previous lactation. The risk and rate of recurrence were always taken to be constant throughout a lactation period and for all the lactations in the selected animals.

The estimations of the parameters obtained with the final model as well as the estimations of their standard deviation and 95% confidence interval are presented in Table I. It can thereby be observed that the risk of REX when the animal is indoors is 7 times greater than when the animal is at pasture. These REX risks no longer depend

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimation</th>
<th>Standard deviation</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of REX indoors</td>
<td>74·10⁻⁵</td>
<td>21·10⁻⁵</td>
<td>(43; 128)·10⁻⁵</td>
</tr>
<tr>
<td>Risk of REX at pasture</td>
<td>11·10⁻⁵</td>
<td>8·10⁻⁵</td>
<td>(3; 45)·10⁻⁵</td>
</tr>
<tr>
<td>Risk of recurrence</td>
<td>1691·10⁻⁵</td>
<td>190·10⁻⁵</td>
<td>(1357; 2106)·10⁻⁵</td>
</tr>
<tr>
<td>Rate of recurrence</td>
<td>0.39</td>
<td>0.03</td>
<td>(0.33; 0.46)</td>
</tr>
<tr>
<td>Infection rate at calving in primiparous cows</td>
<td>0.08</td>
<td>0.03</td>
<td>(0.04; 0.18)</td>
</tr>
<tr>
<td>Infection rate at calving in multiparous cows not having had clinical mastitis in the previous lactation</td>
<td>0.24</td>
<td>0.03</td>
<td>(0.18; 0.31)</td>
</tr>
<tr>
<td>Infection rate at calving in multiparous cows having had clinical mastitis in the previous lactation</td>
<td>0.51</td>
<td>0.05</td>
<td>(0.41; 0.61)</td>
</tr>
</tbody>
</table>
on the lactation stage, the parity number considered or on the calving month.

The risk of recurrence was estimated to be 0.01691 (0.00190) which corresponds to a 95% confidence interval of (0.01357; 0.02106). The mean number of days separating two cases of mastitis, where the second was a recurrence of the first, was therefore estimated to be 1/0.01691 i.e. 59 days or 8.5 weeks, with a 95% confidence interval of (47; 74) days or (6.5; 10.5) weeks. The recurrence rate was estimated to be 0.39 (0.03) which corresponds to a 95% confidence interval of (0.33; 0.46).

The infection rate at calving for the primiparous cows was estimated to be 0.08 (0.03) which corresponds to a 95% confidence interval of (0.04; 0.18). For multiparous cows, this rate depended on the health situation during the previous lactation, and was equal to 0.24 (0.18; 0.31) if the previous lactation was free from clinical mastitis whereas it was equal to 0.51 (0.41; 0.61) if to the contrary, i.e. a rate that was twice as high.

With this final model, estimations of the distribution functions $P(T_1 \leq t)$ and $P(T_2 \leq t)$ presented in Section 2.2 and used in the publication cited above [17], were obtained in continuous lines in Figure 2.

We also obtained distributions of the number of cases of mastitis estimated per lactation by explicitly taking into consideration the relationship between consecutive events within lactation and between consecutive lactations, to be compared with the distribution of the number of mastitis cases observed: these distributions are presented in Table II. It can be observed that the approach chosen makes it possible to adjust the data observed correctly by reducing the deviation from these observed data, both for all the lactations and also for the lactations distributed according to the modalities of the following factors: parity, calving month, breed and productive

**Figure 2.** The distribution function estimates $P(T_1 \leq t)$ and $P(T_2 \leq t)$ obtained with the final model for the two first clinical mastitis with the Orcival data set.
Table II. Distribution of the number of mastitis cases observed and estimated using the final model for the 693 lactations at Orcival.

<table>
<thead>
<tr>
<th>Number of mastitis cases per lactation</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number observed</td>
<td>481</td>
<td>130</td>
<td>45</td>
<td>37</td>
</tr>
<tr>
<td>Number estimated</td>
<td>478</td>
<td>125</td>
<td>54</td>
<td>35</td>
</tr>
</tbody>
</table>

Table III. Distribution of the number of mastitis cases estimated using the final model for the 693 lactations at Orcival, for each modality of the parity factor (lactation number observed in brackets). (5.96 $\chi^2$ value and 0.82 $P$-value for 10 degrees of freedom.)

<table>
<thead>
<tr>
<th>Number of mastitis cases estimated (observed) per lactation</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity 1</td>
<td>201 (202)</td>
<td>27 (30)</td>
<td>18 (15)</td>
<td></td>
</tr>
<tr>
<td>Parity 2</td>
<td>116 (115)</td>
<td>37 (38)</td>
<td>27 (27)</td>
<td></td>
</tr>
<tr>
<td>Parity 3</td>
<td>73 (81)</td>
<td>27 (26)</td>
<td>20 (13)</td>
<td></td>
</tr>
<tr>
<td>Parity 4</td>
<td>44 (41)</td>
<td>16 (19)</td>
<td>11 (11)</td>
<td></td>
</tr>
<tr>
<td>Parity 5 and above</td>
<td>45 (42)</td>
<td>18 (17)</td>
<td>13 (16)</td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Distribution of the number of mastitis cases estimated using the final model for the 693 lactations at Orcival, for each modality of the calving month factor (lactation number observed between brackets). (13.50 $\chi^2$ value and 0.33 $P$-value for 12 degrees of freedom.)

<table>
<thead>
<tr>
<th>Number of mastitis cases estimated (observed) per lactation</th>
<th>0</th>
<th>1</th>
<th>2 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>33 (29)</td>
<td>10 (13)</td>
<td>7 (8)</td>
</tr>
<tr>
<td>November</td>
<td>204 (194)</td>
<td>46 (60)</td>
<td>34 (30)</td>
</tr>
<tr>
<td>December</td>
<td>72 (74)</td>
<td>20 (16)</td>
<td>15 (16)</td>
</tr>
<tr>
<td>January</td>
<td>63 (69)</td>
<td>19 (13)</td>
<td>13 (13)</td>
</tr>
<tr>
<td>February</td>
<td>50 (52)</td>
<td>15 (15)</td>
<td>10 (8)</td>
</tr>
<tr>
<td>March to September</td>
<td>56 (63)</td>
<td>16 (13)</td>
<td>11 (7)</td>
</tr>
</tbody>
</table>

potential (IP – initial production) whereas these factors no longer explicitly intervene in the model. Since these factors had not been found to be significant apart from the influence of the “primiparous/multiparous” factor on the infection rate of the udder at calving, the corresponding distributions are presented respectively in Tables III, IV, V and VI.

3.2. Prediction of the distribution of the number of cases of mastitis per lactation if there is no recurrence, using data from the Orcival site

On the basis of estimations made by the final model obtained previously, it is possible to predict the distribution of the
number of mastitis cases per lactation that should have been observed if there had been no recurrence, i.e. if the treatments applied in the cases of clinical mastitis had been totally effective against the germs present and if all the udders were free of germs at calving. This distribution was obtained by using the estimations of the risks of REX and by setting the REN and the infection rate at calving at a value of zero. This distribution is presented in Table VII and should be compared to the distribution observed. It can thereby be noted on this farm that it should not usually be possible to observe, in a herd of 100 animals during a calving campaign, a lactation with more than one episode of clinical mastitis; the probability of observing lactations with 2 and more cases of clinical mastitis is estimated to be less than 1%.

3.3. Prediction for the animals at the Marcenat site using the previous estimations of the model parameters obtained using data from the Orcival site

On the basis of the risk estimates made using the Orcival data and also with the modalities of the factors of the Marcenat data, the distribution of the number of

Table V. Distribution of the number of mastitis cases estimated using the final model for the 693 lactations at Orcival, for each modality of the breed factor (lactation number observed between brackets). (4.56 $\chi^2$ value and 0.60 $P$-value for 6 degrees of freedom.)

<table>
<thead>
<tr>
<th>Number of mastitis cases estimated (observed) per lactation</th>
<th>0</th>
<th>1</th>
<th>2 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breed MO*PN</td>
<td>173 (174)</td>
<td>47 (49)</td>
<td>33 (30)</td>
</tr>
<tr>
<td>Breed HO</td>
<td>221 (229)</td>
<td>55 (55)</td>
<td>38 (30)</td>
</tr>
<tr>
<td>Breed HO*FF</td>
<td>85 (78)</td>
<td>24 (26)</td>
<td>17 (22)</td>
</tr>
</tbody>
</table>

Table VI. Distribution of the number of mastitis cases estimated using the final model for the 693 lactations at Orcival, for each modality of the IP factor (lactation number observed between brackets). (4.71 $\chi^2$ value and 0.79 $P$-value for 8 degrees of freedom.)

<table>
<thead>
<tr>
<th>Number of mastitis cases estimated (observed) per lactation</th>
<th>0</th>
<th>1</th>
<th>2 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 18 kg</td>
<td>148 (153)</td>
<td>23 (20)</td>
<td>16 (13)</td>
</tr>
<tr>
<td>From 18 to 22 kg</td>
<td>99 (102)</td>
<td>24 (26)</td>
<td>17 (12)</td>
</tr>
<tr>
<td>From 22 to 26 kg</td>
<td>108 (100)</td>
<td>35 (40)</td>
<td>25 (28)</td>
</tr>
<tr>
<td>Over 26 kg</td>
<td>123 (126)</td>
<td>44 (44)</td>
<td>31 (29)</td>
</tr>
</tbody>
</table>

Table VII. Distribution of the number of mastitis cases observed and estimated using the final model without recurrence, for the 693 lactations at Orcival.

<table>
<thead>
<tr>
<th>Number of mastitis cases per lactation</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number observed</td>
<td>481</td>
<td>130</td>
<td>45</td>
<td>37</td>
</tr>
<tr>
<td>Number estimated using the model without recurrence</td>
<td>621</td>
<td>68</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
A recurrent mastitis model in dairy cows 97

This distribution is to be compared with the observed distribution presented in Table VIII. It can thereby be noted that the prediction obtained for this second set of data was accurate, without the need to estimate the model parameters again. The same is true when the lactations are divided according to the modalities of the factors parity, calving date or breed as shown respectively in Tables IX, X and XI.

### 3.4. Prediction for the animals at the Rennes site using the previous estimations of the model parameters obtained using data from the Orcival site

On the basis of the risk estimations made using the Orcival data and also with the modalities of the factors of the Rennes data, the distribution of the number of mastitis cases per lactation was predicted.
This distribution is to be compared with the observed distribution presented in Table XII. It can thereby be noted that the prediction obtained for this second set of data was accurate, without the need to estimate the model parameters again. The same is true when the lactations are divided according to the modalities of the factors calving month and productive potential (IP – initial production) as shown in Tables XIII and XIV. However, the predictions were less accurate for parity factor (Tab. XV), mainly for the primiparous cows having contracted clinical mastitis only once during lactation.

4. DISCUSSION

4.1. Model validation

This new model made it possible not only to model the distribution of the number of mastitis cases per lactation from the construction data from the Orcival site (Tab. II), but also from the validation data
A recurrent mastitis model in dairy cows

with all the lactations in the Marcenat (Tab. VIII) and Rennes (Tab. XII) sites over the different periods indicated, for sites in two French regions that are very distinct in geographical terms. These good results of prediction were also obtained when the results were finely distributed according to the modalities of important potential factors (“risk factor” retained in previous studies with classical analysis) not present in the final model: parity (Tabs. III and IX), calving month (Tabs. IV, X and XIII), breed (Tabs. V and XI) and IP (Tabs. VI and XIV) factors.

### 4.2. Biological model interpretation

The approach chosen still consisted of describing the relationship between consecutive cases of clinical mastitis using differentiated parameters expressing notions of clinical mastitis due to exogenous infection (REX), and of clinical mastitis due to persistent endogenous infection (REN). This approach makes it possible to take into consideration the relative effectiveness of antibiotic treatments applied during lactation following an episode of mastitis [32, 42] or during the drying off period, treatments that do not systematically ensure that the udder will recover bacteriologically [23, 38]. The new idea introduced to take into account the possible relationship between consecutive lactations is simply based on assessing the health status of the udder at calving: “infected” or “not infected”. The approach chosen closely corresponds to the current observations of veterinary epidemiologists that are convinced that, in order to improve the knowledge and assessment of the pathology “clinical mastitis” in farms, it is necessary to explicitly take into consideration these relationships within lactation and between

<table>
<thead>
<tr>
<th>Number of mastitis cases estimated (observed) per lactation</th>
<th>0</th>
<th>1</th>
<th>2 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 18 kg</td>
<td>51 (47)</td>
<td>10 (13)</td>
<td>7 (8)</td>
</tr>
<tr>
<td>From 18 to 22 kg</td>
<td>68 (64)</td>
<td>12 (20)</td>
<td>8 (5)</td>
</tr>
<tr>
<td>From 22 to 26 kg</td>
<td>43 (41)</td>
<td>14 (20)</td>
<td>10 (7)</td>
</tr>
<tr>
<td>Over 26 kg</td>
<td>92 (88)</td>
<td>38 (35)</td>
<td>28 (35)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of mastitis cases estimated (observed) per lactation</th>
<th>0</th>
<th>1</th>
<th>2 and above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity 1</td>
<td>121 (112)</td>
<td>19 (30)</td>
<td>13 (11)</td>
</tr>
<tr>
<td>Parity 2</td>
<td>59 (62)</td>
<td>24 (29)</td>
<td>18 (9)</td>
</tr>
<tr>
<td>Parity 3</td>
<td>35 (31)</td>
<td>14 (12)</td>
<td>10 (16)</td>
</tr>
<tr>
<td>Parity 4 and above</td>
<td>40 (35)</td>
<td>18 (17)</td>
<td>14 (19)</td>
</tr>
</tbody>
</table>
consecutive lactations in assessment and prediction models [23]. This modelling approach is based solely on parameters that can be estimated and easily interpreted in epidemiological terms (Fig. 1). With 0.01691 (Tab. I) for recurrence risk (REN) mean estimation per day (λ), the time-lapse mean estimation (1/λ) for a new case of clinical mastitis is around 59 days (or 41 days for time-lapse median estimation, with a theoretical median formula ln(2)/λ). This value is in the day interval used by a different author: 8 to 90 days. This high value is also biologically compatible with the infection persistence in lactation. For many cows having infection at calving, it is also logical to observe a clinical mastitis especially in the first month of lactation (50% of recurrence of clinical mastitis in the first 41 days after calving).

A model is obtained that provides predictions for the distribution of the number of events per lactation (Tabs. II, VIII and XII), for the distribution of instances of the occurrence of consecutive events (Fig. 2), but also for the distribution of the number of events per lactation in lack of REN parameters in a herd (Tab. VII).

4.3. Confounding factors on REX parameter

One of the significant consequences associated with integrating the environmental factor “animal indoors or at pasture” and the health status of the udder at calving into the model, is the disappearance of the factors lactation stage and calving month. The REX risk remains constant when the animal remains indoors and does not decrease thereafter to another constant level when the animal goes out to pasture. The higher risk of REX at the beginning of lactation observed by a majority of authors would therefore simply be the result of the fact that the following were not taken into consideration in their models: the health status of the udder at calving and the environment surrounding the animal during production: “indoors” or “at pasture”, with the risk of REX for an animal indoors always being considered to be higher than the risk of REX of an animal at pasture. For the same reasons, the potential effects of parity and calving month are no longer significant vis à vis the risk of REX. This shows how important the chosen model and its defining parameters are in demonstrating and identifying real factors influencing the “clinical mastitis” phenomenon. This also shows how crucial the validation stage is, a stage that should follow on from any modelling stage and which consists of applying the model developed to new data without re-estimating the parameter values (Tabs. VIII to XV).

4.4. Initial production factor and “infection” rate at calving

On the contrary, for the Rennes site, it can be noted that the model underestimated the number of lactations in primiparous cows that had not been subject to a single case of clinical mastitis during lactation (Tab. XV). For this same site, the number of lactations with initial production of below 22 kg and free of clinical mastitis during the lactation was also underestimated (Tab. XIV). Yet in this site, 88% of primiparous cows had an initial production of below 22 kg (comparable to the value of 91% at Orcival), whereas 90% of multiparous cows initially produced over 22 kg (a value that is higher than the 77% for Orcival). If we are to adhere to what different authors appear to have observed (increased risk in multiparous cows with a higher initial production), we should have obtained a bias for the multiparous cows for which the initial production was above 22 kg, which was not the case here. Since the estimation bias concerned lactations with initial production of below 22 kg, this mainly concerned primiparous cows. This bias observed in the initial production for lactations having been subject to just one case of clinical mastitis during lactation only serves to accurately reflect the bias.
observed for primiparous cows also having been subject to just one case of clinical mastitis during lactation. At Rennes, the percentage of primiparous cows (54%) was much higher than at Orcival (36%) or Marcenat (29%), just as it is only at the Rennes site and for primiparous cows that it is possible to observe a lactation percentage (37%) lasting from the calving date to the day on which the animal went back indoors, ranging between 30 and 60 days, a percentage that is higher than at Orcival (12%) or Marcenat (28%). “The incidence of clinical mastitis was lower on farms when the average dry period was less than 40 d” for example, was observed by Peeler et al. in Great Britain [33]. The fact that only the number of lactations having been subject to just one case of clinical mastitis is biased indicates that, for the primiparous cows at the Rennes site, the “infection” rate at calving is underestimated and that this rate may depend on the length of time spent indoors before calving. This means that the longer the time spent indoors, the higher the “infection” rate at calving. This phenomenon has already been observed in multiparous cows that have an “infection” rate at calving estimated to be higher than for primiparous cows and spend a greater amount of time indoors before calving (28% of multiparous cows spend less than 30 days indoors before calving at Orcival, compared to 64% of primiparous cows). In multiparous cows, no considerable bias was observed for either of the three sites because the lactation percentage values corresponding to indoor times of below 30 days, between 30 and 60 days or above 60 days, are plainly comparable. These observations would therefore tend to indicate that part of the “infection rate at calving” parameter might be explained by infections contracted during the time spent indoors and having remained uncured during the dry period before calving in multiparous cows, or in the two months before calving for example in primiparous cows.

The period of two months was chosen for primiparous cows so as to be similar to the mean dry period of multiparous cows.

4.5. A potential model improvement

A simple way of making this model evolve, allowing it to take into consideration this phenomenon of infections contracted and remaining uncured during the period before calving, would be to divide parameter $q$ corresponding to the “infection rate at calving”. An example of a suitable way to divide it up would be as follows: $q = 1 - [\{(1 - q_{\alpha}) \cdot \exp(-\mu_{\text{out}} \cdot D_{\text{out}}) \cdot \exp(-\mu_{\text{in}} \cdot D_{\text{in}})\}]$, where “$q_{\alpha}$” is the infection rate just after drying off and applying possible systematic treatment during drying off, “$D_{\text{out}}$” and “$D_{\text{in}}$” are the pre-calving periods respectively at pasture and indoors, and “$\mu_{\text{out}}$” and “$\mu_{\text{in}}$” are the risks of contracting an infection that will not be cured by the following calving. This would amount to explaining parameter $q$ by the time spent indoors and the time spent at pasture before calving, depending on the parameters relating to the corresponding risk of infection and a chronic infection rate just after drying off. It appears to be just as obvious that the significant effect demonstrated in this study of the factor “having had or not had clinical mastitis in the previous lactation” on parameter $q$ would naturally affect parameter “$q_{\alpha}$”. It would only remain to complete the explanation of this parameter with a factor or co-variable pertinent characterising the fact that a cow has, in the previous lactation, been subject to a state of “chronic” infection yet without having had clinical mastitis.

Identifying a pertinent indicator of the “chronic infection” status of an animal’s udder during production in the previous lactation and integrating it into the model as a co-variable or factor influencing parameter “$q_{\alpha}$” would certainly make it possible to further increase the pertinence of the model proposed, while at the same time providing an unbiased estimate of the
part of parameter “q” due to infections contracted and not spontaneously cured during the dry period before calving. A potential candidate for this rank of pertinent indicator could, for example, be the mean cell count level calculated on the basis of a few monthly cell counts from the previous lactation, cell counts chosen from the last of the lactation and counts carried out by the French departmental Milk Control body, for example. Since these count values were not available in the experimental farms, we were unable to test its potential utility. One of the next objectives of our work will therefore be to assess the advantages of such an indicator or of any other pertinent indicator making it possible to characterise the health status of a “non-infected” or “chronically infected” udder during a given period and that is available in most French farms, in order to test its usefulness in the model proposed. In doing so, we will of course come face to face with the difficulties encountered by most of the authors that have studied the use of cell counts to assess the health status of an animal’s udder, and in particular to assess its possible sub-clinical scope. With respect to the risk of infection not cured spontaneously during the dry period, one possible way of going deeper into the matter would, for example, be to take into account the time spent at pasture and indoors during this dry period; it appears to be relatively obvious that the risk would certainly be higher if the animal were indoors, as is the case during production.

4.6. Prospect

All that then remains to be done is to study the “usual” potential farming system factors that may influence the REX risk level indoors or at pasture, or even the risk of contracting a chronic infection during the dry period, indoors or at pasture. This of course could only be done on the basis of a considerable number of farms in which it is possible to make continuous and individual recordings (per animal) of the occurrence of clinical mastitis over a period of at least one complete lactation per animal. This type of recording combined with a relatively precise survey into breeding practices is currently underway in approximately 600 French farms representative of different breeding practices used in the country.

ACKNOWLEDGEMENTS

We would like to thank those who participated in the data collection (those working in and in charge of the experimental farms: Jean-Paul Garel and Philippe Pradel for the Marcenat site, Jacques Bony and Dominique Pomiès for the Orcival site and Jean-Louis Trocon for the Rennes site) without whom the elements necessary for developing and validating the model proposed would not have been available to us. Also we would like to thank Anne-Marie Wall, of INRA Jouy-en-Josas linguistic service, for the translation of this article.

REFERENCES

[6] Barnouin J., Chassagne M., Factors associated with clinical mastitis incidence in...
A recurrent mastitis model in dairy cows


[32] Owens W.E., Ray C.H., Watts J.L., Yancey R.J., Comparison of Success of Antibiotic Therapy During Lactation and Results of


