The Norwegian Healthier Goats program—Modeling lactation curves using a multilevel cubic spline regression model

G. E. Nagel-Alne,*† R. Krontveit,‡ J. Bohlin,§ P. S. Valle,¶ E. Skjerve,† and L. S. Solvørød*  
*TINE Norwegian Dairies, PO Box 58, NO-1431 As, Norway  
†Department of Production Animal Clinical Science, and  
‡Department of Food Safety and Infection Biology, Norwegian School of Veterinary Science, PO Box 8146 Dep., NO-0033 Oslo, Norway  
§Norwegian Institute of Public Health, Division of Epidemiology, Marcus Thrane gate 6, PO Box 4404, 0403 Oslo, Norway  
# Kontali Analyse AS, Industriveien 18, 6517 Kristiansund, Norway

ABSTRACT

In 2001, the Norwegian Goat Health Service initiated the Healthier Goats program (HG), with the aim of eradicating caprine arthritis encephalitis, caseous lymphadenitis, and Johne’s disease (caprine paratuberculosis) in Norwegian goat herds. The aim of the present study was to explore how control and eradication of the above-mentioned diseases by enrolling in HG affected milk yield by comparison with herds not enrolled in HG. Lactation curves were modeled using a multilevel cubic spline regression model where farm, goat, and lactation were included as random effect parameters. The data material contained 135,446 registrations of daily milk yield from 28,829 lactations in 43 herds. The multilevel cubic spline regression model was applied to 4 categories of data: enrolled early, control early, enrolled late, and control late. For enrolled herds, the early and late notations refer to the situation before and after enrolling in HG; for nonenrolled herds (controls), they refer to development over time, independent of HG. Total milk yield increased in the enrolled herds after eradication: the total milk yields in the fourth lactation were 634.2 and 873.3 kg in enrolled early and enrolled late herds, respectively, and 613.2 and 701.4 kg in the control early and control late herds, respectively. Day of peak yield differed between enrolled and control herds. The day of peak yield came on d 6 of lactation for the control early category for parities 2, 3, and 4, indicating an inability of the goats to further increase their milk yield from the initial level. For enrolled herds, on the other hand, peak yield came between d 49 and 56, indicating a gradual increase in milk yield after kidding. Our results indicate that enrollment in the HG disease eradication program improved the milk yield of dairy goats considerably, and that the multilevel cubic spline regression was a suitable model for exploring effects of disease control and eradication on milk yield.

Key words: lactation curve, goat, cubic splines, eradication program

INTRODUCTION

The chronic and infectious diseases caprine arthritis encephalitis (CAE), Johne’s disease (caprine paratuberculosis), and caseous lymphadenitis (CL) are believed to influence milk production and thus lactation curves. As a consequence, the Norwegian Healthier Goats eradication and control program, Healthier Goats (HG) for short, was initiated in 2001 with the aim of eradicating CAE, CL, and caprine paratuberculosis from Norwegian goat herds.

Caprine arthritis encephalitis is caused by the CAE virus (CAEV). Nord et al. (1998) reported that almost 90% of Norwegian dairy goat herds had antibodies against CAEV. Common clinical manifestations are chronic proliferative synovitis and periarteritis, interstitial pneumonia, indurative mastitis (“hard udder”), and chronic wasting in adult goats. Acute afebrile leukoencephalomyelitis is characteristic of the infection in goat kids. Does infected with CAEV often have a viral infection of the mammary gland, which is a target organ for the virus (Kennedy-Stoskopf et al., 1985; Lerondelle et al., 1995). Peterhans et al. (2004) estimated milk production losses to be approximately 10%. Lesions in the udder can occur before puberty and gradually become more extensive (Dawson, 1987). Quantitative data relating to the financial impact of CAEV infection are limited, but some studies suggest that indurative mastitis is the most economically significant component of CAEV through reduction of milk yield, most pronounced in older does (Krieg and Peterhans, 1990; Greenwood, 1995; Martínez-Navalón et al., 2013). The disease is mainly spread through the ingestion of virus-infected goat colostrum or milk by kids and indirectly via the environment (Adams et al., 1983).
Caprine paratuberculosis is caused by *Mycobacterium avium* ssp. *paratuberculosis*. The fecal–oral route is important in disease transmission, and chronic, progressive weight loss is the most common clinical manifestation in adult goats (Manning and Collins, 2001). In Norway, caprine paratuberculosis has been endemic in the southern part of the country for decades (Djønne, 2003). A temporary vaccination program reduced the infection rate substantially (Saxegaard and Fodstad, 1985), but replication and shedding of the bacteria continued (Djønne, 2003). Several reports describe the economic losses due to decreased milk production in dairy cows with paratuberculosis (Stabel, 1998; Radostits et al., 2007). Similar losses would be expected with caprine paratuberculosis, but no specific findings on the economic impact have been reported (Smith and Sherman, 2009).

Caseous lymphadenitis is caused by *Corynebacterium pseudotuberculosis* and is a chronic disease that mainly affects sheep and goats (Brown and Olander, 1987; Williamson, 2001). The bacterium spreads through wounds or small breaks in the skin and localizes in regional lymph nodes, resulting in local superficial abscesses, commonly appearing after 2 to 6 mo (Ashfaq and Campbell, 1980; Kuria et al., 2001). Internal abscess formation can occur if the bacteria reach the thoracic lymph duct or are inhaled (Smith and Sherman, 2009). The superficial abscesses may rupture and drain spontaneously, which leads to contamination of both the environment and other individuals in the herd (Ashfaq and Campbell, 1979a,b; Holstad, 1986). Abscess formation can also occur in the mammary lymph nodes (Kuria et al., 2001). In sheep, CL can induce suppurative mastitis with impaired milk production, and a similar condition is likely to occur in goats (Baird and Fontaine, 2008).

As a result of the control and eradication in HG, the status of CAE, caprine paratuberculosis, and CL in Norwegian dairy goats has changed from enzootic to rare and close to eradicated with the inherent potential for improved milk yield and milk quality.

The graphical representation of daily milk yield against time after parturition is called a lactation curve (Sherchand et al., 1995). Modeling daily milk yield by use of lactation curves is a useful management tool in goat milk production (Gipson and Grossman, 1990; Ruiz et al., 2000) and can be used to estimate the animals' requirement for diet formulations throughout the different phases of lactation (León et al., 2012). Also, early detection of animals with low production can identify underlying pathological factors and thus allow adequate treatment or culling (Dematawewa et al., 2007).

The mathematical functions applied for the presentation of lactation curves should allow a valid reflection of the milk production throughout the lactation period (León et al., 2012). The geometric shape (time of peak yield and persistency) and scale (initial and peak yield) parameters are important characteristics of a lactation curve. Breed, kidding season, age of doe, litter size, and herd and feeding practices are important factors influencing the lactation curve shape (Gipson and Grossman, 1990; Motaldo et al., 1997; Macciotta et al., 2005).

Mathematical functions used to model lactation curves; for example, Wood (1967), estimate fixed coefficients that describe the shape and scale of the lactation curve. Such models can be hampered by the bias introduced by the mathematical formula used, resulting in reduced flexibility compared with other statistical methods specifically constructed for data adaptation. By applying data-driven, mixed-effects multilevel regression models, it is possible to study how both fixed and random effects influence the outcome (Dohoo et al., 2009). Random predictor variables can additionally be used to model variation at several hierarchical levels such as farm, individual animal, and parity, as predicted by the parameters included in the model. By fitting functions such as splines at certain points (knots), a flexible curve can be estimated to express test-day milk recordings (White et al., 1999).

The aim of the study presented here was to model lactation curves using a combination of the methods described above. In the present study, lactation curves from HG-enrolled and nonenrolled herds were examined and compared using a multilevel cubic spline regression model to study the effect of HG on milk yield in Norwegian dairy goat herds.

**MATERIALS AND METHODS**

**HG Control and Eradication Program**

The HG program was initiated as a voluntary control and eradication program for Norwegian goat herds. The farmers had costs related to the eradication of CAE, CL, and caprine paratuberculosis reimbursed by HG. In 2012, a reduced milk price was introduced for farmers not yet enrolled in HG, changing the decision criteria for farmers regarding enrollment, or not, in HG. The incentive caused by the reduced milk price enabled enrollment of all Norwegian dairy goat herds in HG by the end of 2012 (Nagel-Alne et al., 2014).

The “snatching” method applied in HG is based on the principle of raising infection-free kids and culling all does at the end of lactation. The kids are snatched...
from the birth canal at delivery and raised in a clean compartment using cow colostrum or milk replacer. The snatching method was used on farms on which at least 1 of the 3 diseases (CAE, CL, or caprine paratuberculosis) was clinically present. Herds without CL and caprine paratuberculosis infection and with less than 10% CAE seropositive goats were allowed to cull the test-positive animals and continue production as before. This applied to 18% of the goat milk herds enrolled in HG by the end of 2012. Farmers did not know their herd prevalence of CAE, CL, and caprine paratuberculosis before they decided to enroll in HG. Identification of infection status in the herds was done after the farmers decided to enroll. The material in the present study included HG-enrolled herds that applied the snatching method, meaning that the herds included were those most severely affected by the 3 diseases. Furthermore, the nonenrolled herds were a mix of herds moderately or heavily affected by CAE, CL, and caprine paratuberculosis.

**Target and Source Population**

The target population was all Norwegian dairy goats in dairy herds that were members of the TINE Goat Milk Recording System (GMRS), which collects and archives milk recordings according to the method described in the National Goat Control Handbook (https://medlem.tine.no/trm/tp/binary?id = 15624). The participation rate for dairy goat producers in GMRS was 87.7% in 2012. The source population was selected based on the criteria of either being enrolled in HG or not.

The data on milk recordings used in this study were extracted from the GMRS and included data from 1999 to 2008. The data were separated into 2 categories depending on HG status: enrolled herds (snatching method) and nonenrolled herds (hereafter called control herds).

A regular lactation in Norwegian dairy goats lasts from d 6 to d 275 after kidding. The following inclusion criteria were applied: (1) a minimum of 3 milk yield records in the GMRS per lactation, (2) the first milk yield record within 60 d after kidding, (3) herds must have enrolled in HG from 2002 through 2008, and (4) herds in the control category should not have enrolled in HG any earlier than 2009. Milk yields recorded before d 6 or after d 275 in lactation and milk records <0.2 kg were excluded from the data set as recommended by International Committee of Animal Recording (ICAR, 2012). The mean interval between consecutive milk recordings was 40 to 50 d.

To compare milk yield in enrolled herds with that of control herds while controlling for changes in milk yield over time due to factors other than HG, the data were further categorized along a timeline. Every milk yield recording from the enrolled category was coded according to time of recording in relation to year of HG enrollment. The notation “early” was used for records from before HG enrollment (years 1999 to 2005) and “late” for records after HG enrollment (years 2006 to 2008). Depending on the actual year of enrolling in HG, different years were coded as early or late in different herds. Similarly, in the control herds, records from years 1999 to 2003 were considered as “early” and records from the years 2004 to 2008 were considered as “late.” This classification divided the data into 4 categories: enrolled early, control early, enrolled late, and control late. The original data set included 247,617 records, from which 135,446 test-day records corresponding to 28,829 lactations of 9,791 does in 43 herds were extracted. The number of records in the 4 different categories is outlined in Table 1.

**Statistical Analysis**

The overall approach was to build a model that incorporated both fixed and random effect parameters to explore their effect on the resulting milk yield. The statistical analyses were conducted in Stata SE/11.2 (Stata Corp., College Station, TX) using the xtmixed procedure for multilevel mixed-effects regression models. A multilevel regression model was fit on the whole data set and expressed as

\[
y_{ijkl} = a_0 + f(x_{ijkl}) + w_k + Z_{ijkl}x_{ijkl} + e_{ijkl}, \tag{1}
\]

where \(y_{ijkl}\) represents the output as recorded daily milk yield, \(a_0\) is the intercept, \(x\) and \(w\) are DIM and parity, as the fixed and explanatory variables, respectively, and \(f(x)\) is the function of DIM. Parity was categorized from 1 to 5, where the latter category included lactations 5 to 16. The random effects parameters in the model were farm \((i)\), individual goat \((j)\), and lactation number \((k)\).
(k), giving the model a 3-level hierarchy to control for variability between herds, the individual goat, and lactation number. Each observation is designated by l, and \( e_{ijkl} \) represents the corresponding residual errors. The \( Z_{ijk} \) is the matrix representing the hierarchical random effects for level i, j, k. The model was extended with a random slope to allow for a nonconstant variation of \( x_i \) with respect to lactation ID.

The functional relationship between the outcome (milk yield on test-day) and DIM was evaluated by exploring several different transformations \( f(\cdot) \) of DIM in the overall model explained above in Equation [1], including linear spline, cubic spline, and polynomial transformation of DIM. For the linear and cubic spline transformations of DIM, knots were placed at 5 percentiles; namely 49, 78, 112, 157, and 210 DIM, as recommended by Harrell (2001). The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to compare and rank the models (Akaike, 1974; Schwarz, 1978; Raftery, 1995). A reduction in AIC and BIC of 5 and between 6 and 10 units, respectively, was considered a significant improvement in AIC and BIC of 5 and between 6 and 10 units, with respect to lactation ID.

The cubic spline regression model was expressed as

\[
Y_{ijl} = a_0 + f(x_{ijkl}) + Z_{ijkl} + e_{ijkl}
\]

where \( a_0 \) was the intercept, \( x \) was the cubic spline-transformed DIM being the only fixed effect, and where farm \((i)\) and individual goat \((j)\) were the random effects; \( Z_{ijkl} \) was the matrix representing the hierarchical random effects for level i, j. Each observation is designated by l, and \( e_{ijkl} \) is the residual error at each level.

The cubic spline regression model was applied to the 4 subsets; namely: enrolled early, control early, enrolled late, and control late, which again were further separated into parities 1 to 4 with the aim to estimate the effect of HG on milk yield in each parity. The number of parities was reduced to 4 due to sparse records in higher lactations, thus parity 4 included lactations 4 to 16. The peak yield (PY) and day of peak yield (DPY) were calculated as the maximum value from estimated parameter values, whereas the total yield (TY) was calculated from the integral of the curves (area under the curve) for each of the parities 1 to 4 in the 4 categories. Through the scatterplot of point estimates, lowess smoothing curves (Cleveland, 1979) were applied to generate lactation curves for the 4 categories. The significance of the random effects of herd, individual goat, and lactation ID were evaluated through a likelihood ratio test (LRT) of models with and without each random effect, where a one-sided test \( P < 0.05 \) was considered significant. The proportion of variance at each random effect level was calculated by relating the variance at the specific level to the total variance.

**RESULTS**

The results of fitting the multilevel cubic spline regression model (2) to the 4 categories of data (enrolled early, control early, enrolled late, and control late) separated by parity (1 to 4) are summarized in Table 3, and the accompanying lactation curves for the different parities in the 4 categories are visualized in Figure 1. Total yield increased with parity in all 4 categories (Table 3).

As can be seen from Table 3, in all 4 parity groups, milk production increased between the early and the late periods but the increase was greater in the enrolled groups than in the control groups. Further, the TY for parity 4 increased by 239 kg from the enrolled early to the enrolled late category, whereas the corresponding increase in TY in the categories control early to control late was 88 kg. In other words, TY after HG enrollment was 63% higher compared with TY for the control late category. In the 2 control categories, the highest TY was reached in parity 3, whereas in the 2 enrolled categories, the highest TY was found in parity 4.

The DPY ranged from 49 to 56 in different parities in the enrolled late category. For the control categories, the DPY is estimated to occur much earlier (d 6 in lactation) for parities 2, 3, and 4. A slight increase in the slope toward PY can be seen for all parities in the enrolled late category and corresponds well to the reported DPY in Table 3. For both the control early and control late categories, the lactation curves showed a steady decline from the start toward the end of the lactation period, with no actual peak yield being evident. For the enrolled early category, the first lactation shows an increase toward PY, whereas the later lactations have a similar pattern as the control categories.

The proportion of variance was estimated for the 4 categories, with parity included as a fixed effect in each category; that is, the variance components include all

<table>
<thead>
<tr>
<th>Transformation of DIM</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second-degree polynomial</td>
<td>230553.3</td>
<td>230661.3</td>
</tr>
<tr>
<td>Third-degree polynomial</td>
<td>230662.8</td>
<td>230770.8</td>
</tr>
<tr>
<td>Linear spline</td>
<td>229513.7</td>
<td>229661.0</td>
</tr>
<tr>
<td>Cubic spline</td>
<td>229503.7</td>
<td>229641.2</td>
</tr>
</tbody>
</table>

Table 2. Akaike information criteria (AIC) and Bayesian information criteria (BIC) for different transformations of DIM in an overall regression model for prediction of lactation curves in Norwegian dairy goats.
Table 3. Total milk yield (kg of milk/lactation), peak milk yield (maximum kg of milk/d), day of peak milk yield and the proportion of variance at herd, goat, and lactation number levels for the categories enrolled early, enrolled late, control early, and control late in parity 1, 2, 3, ≥4 estimated from a cubic spline regression model in Norwegian dairy goats.

<table>
<thead>
<tr>
<th>Categories with parities 1 to ≥4 (no. of records)</th>
<th>Total milk yield</th>
<th>Peak milk yield</th>
<th>Day of peak milk yield</th>
<th>Proportion of variance¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrolled early</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (11,115)</td>
<td>496.2</td>
<td>2.07</td>
<td>58</td>
<td>0.18</td>
</tr>
<tr>
<td>2 (9,365)</td>
<td>592.5</td>
<td>2.47</td>
<td>39</td>
<td>0.21</td>
</tr>
<tr>
<td>3 (6,828)</td>
<td>632.6</td>
<td>2.65</td>
<td>6</td>
<td>0.21</td>
</tr>
<tr>
<td>≥4 (12,704)</td>
<td>634.2</td>
<td>2.64</td>
<td>6</td>
<td>0.21</td>
</tr>
<tr>
<td>Enrolled late</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (15,348)</td>
<td>621.1</td>
<td>2.63</td>
<td>56</td>
<td>0.16</td>
</tr>
<tr>
<td>2 (10,706)</td>
<td>787</td>
<td>3.31</td>
<td>49</td>
<td>0.25</td>
</tr>
<tr>
<td>3 (5,631)</td>
<td>858.6</td>
<td>3.65</td>
<td>52</td>
<td>0.16</td>
</tr>
<tr>
<td>≥4 (4,568)</td>
<td>873.3</td>
<td>3.74</td>
<td>54</td>
<td>0.19</td>
</tr>
<tr>
<td>Control early</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (7,342)</td>
<td>518.2</td>
<td>2.23</td>
<td>46</td>
<td>0.23</td>
</tr>
<tr>
<td>2 (5,749)</td>
<td>604.8</td>
<td>2.67</td>
<td>6</td>
<td>0.21</td>
</tr>
<tr>
<td>3 (4,391)</td>
<td>636.8</td>
<td>2.98</td>
<td>6</td>
<td>0.21</td>
</tr>
<tr>
<td>≥4 (7,481)</td>
<td>613.2</td>
<td>2.9</td>
<td>6</td>
<td>0.21</td>
</tr>
<tr>
<td>Control late</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (10,451)</td>
<td>555.8</td>
<td>2.35</td>
<td>56</td>
<td>0.23</td>
</tr>
<tr>
<td>2 (8,064)</td>
<td>678.3</td>
<td>3.02</td>
<td>6</td>
<td>0.21</td>
</tr>
<tr>
<td>3 (5,885)</td>
<td>715.9</td>
<td>3.19</td>
<td>6</td>
<td>0.21</td>
</tr>
<tr>
<td>≥4 (9,818)</td>
<td>701.4</td>
<td>3.22</td>
<td>6</td>
<td>0.21</td>
</tr>
</tbody>
</table>

¹Parity was included as categorical fixed effect and proportion of variance includes all parities in each category.

Figure 1. Cubic spline function fitted to lactations recorded for the 4 parities: (a) parity 1, (b) parity 2, (c) parity 3, and (d) parity ≥4 separated into 4 categories of milk yield recordings: enrolled early, enrolled late, control early, and control late in Norwegian dairy goats.
parities in each category. The proportion of variance was high and statistically significant ($P < 0.001$) at each level (herd, individual goat, and lactation number) although it was highest at the individual goat level, ranging from 21 to 25% depending on category (Table 3).

**DISCUSSION**

The main findings in the present study were the substantial variation in shape and scale of the lactation curves between enrolled and control categories, and that PY, DPY, and TY differed between these categories and between parities. Clustering was high and significant at the levels of herd, individual goat, and parity.

Data were categorized into enrolled and nonenrolled and further categorized according to time of recordings related to year of enrolling in HG. This approach was deemed necessary to compare milk yield performance in enrolled and nonenrolled herds, and to measure the effect of eradication on milk yield by controlling for the general development in milk yield performance over time, independent of HG.

The TY after enrolling in HG was significantly higher than that before enrolling. Estimated TY increased by almost 20% when the TY of the fourth parity of enrolled late was compared with that of the fourth parity of the control late category. The diseases eradicated through HG may all impair milk production (Burrell, 1981; Smith and Cutlip, 1988), and disease eradication was probably the main reason behind the increased TY in enrolled herds. In addition, we observed obvious differences in the shape of the lactation curves between enrolled and control herds. It was only in the enrolled late category that an increase in milk production toward a PY at around 50 d after kidding was evident. A possible explanation for the very early DPY (6 DIM) found in the control and enrolled early categories could be that these goats were unable to increase milk production after kidding due to the detrimental effects of the present diseases. Indeed, CAE is known to impair milk yield performance in higher lactations and to cause “hard udder” early after parturition (Robinson and Ellis, 1986; Greenwood, 1995; Martínez-Navalón et al., 2013).

Several other factors might explain the increase in milk yield after disease eradication. In the present study, the highest level of clustering was found at the individual goat level; the proportion of variance was >20%. Information about breeding regimen and selection of breeding animals or breeding values was not available in this study, but enrolled and control herds did have the same breeding system (data not shown). Herd-level clustering was higher in the control late category than in the enrolled late category, 23 and 16%, respectively. This could indicate that the enrolled herds were more homogeneous due to improved management and feeding regimens after disease eradication. Thus, these factors might also contribute to improvement in milk yield. Motivation factors for farmers to join HG have not been screened and no potential systematic difference between early enrolers and late enrolers have been detected or incorporated in this study. Because eradication of the 3 diseases through HG is labor-intensive and requires substantial effort from farmers, we might expect that the most motivated and best organized farmers enrolled first. These are not necessarily the farms with the highest prevalence of the 3 diseases. Milk production was higher in the control early than in the enrolled early category in this study, indicating a higher production level in the control early category before enrolling in HG. The control late category provided information on the trend of milk production independent of HG. The enrolled early and control early categories were considered equal in regard to variation in genetic material, management practice, and prevalence of diseases. Assuming that the controls and enrolled farms represent the same situation before HG enrollment, the comparison of controls and enrolled farms after HG would indicate the effect of disease eradication (and other herd effects due to HG enrollment, as discussed above). The effect of improved management and feeding regimens on daily milk yield can be regarded as independent from the effect of freedom from disease, although the eradication of CAE, CL, and caprine paratuberculosis could affect the goats’ ability to utilize feed optimally.

The intercept and slope of the lactation curve provide important information about aspects related to milk production, such as the level of maximum yield (peak) and the time of peak yield in the lactation period. This enables farmers to plan a stable production level throughout the year, and to organize the feeding of groups according to the different physiological needs determined by the level of milk production.

In the present study, we modeled lactation curves using several different statistical regression models with different transformations of DIM, and found that a multilevel cubic spline-based regression model fitted the data significantly better than other models. Although polynomial regressions can fit data exhibiting nonlinear trends, polynomials can introduce bias because they tend to shoot up or down at end-points and are thus influenced by records at the end-points.

Linear splines allow for the estimation of the relationship between daily milk yield and DIM as a piecewise linear function that connects through the knots. It has

---

*Journal of Dairy Science Vol. 97 No. 7, 2014*
been shown that a linear spline can be used to fit many functions well (Gould, 1993; Panis, 1994; Harrell, 2001). However, cubic splines may be a better choice than linear splines when nonlinear curves, such as lactation curves, are fitted (White et al., 1999).

In recent years, multilevel spline regression models have been introduced to describe milk yield throughout lactation in dairy cattle (Silvestre et al., 2006; Bohmanova et al., 2008; Schaeffer and Jamrozik, 2008). The use of spline functions in the study of lactation curves from dairy goats is more limited, but León et al. (2012) compared the use of a quadratic spline function with several parametric models in dairy goats and found that the quadratic spline function showed the best fit.

Nonlinear regression models, such as Wood (1967), may be superior in terms of predictive abilities, because the statistical nature of such models allows for statistical inference. However, in a purely explorative study, not only are such models complex, but they can be difficult to run if a self-starter routine is absent or fails to estimate the starting values required for the estimation of such models (Pinheiro and Bates, 2000). With the use of multilevel cubic spline regression model in the present study, the traditional shape of a lactation curve, with an ascending and descending phase, was not obvious. A possible reason for this is that the statistical constraints imposed on the curve-fitting procedure by the traditional Wood model are absent. Hence, the traditional lactation curve models, such as Wood (1967), might introduce bias and might be better suited for predictive modeling rather than data exploration, which was the main focus of the present study.

CONCLUSIONS

Enrolling in HG improved the milk yield performance of dairy goats significantly. The multilevel cubic spline regression model was found to be an effective method to model lactation curves. Even though every goat population and country is unique, the application of a general statistical model that is largely data-driven, such as the one used in the present study, probably elucidates the actual shape and scale of the lactation pattern adequately. Therefore, we recommend spline regression models with appropriate random effects included in data-driven lactation curve studies such as that presented here.

ACKNOWLEDGMENTS

The authors thank the funders of this study: TINE Norwegian Dairies (Ås, Norway) and The Norwegian Goat Health Services (TINE Norwegian Dairies). and health data was given by The Norwegian Dairy Goat Recording System (TINE Norwegian Dairies, Ås, Norway) and The Norwegian Goat Health Services (TINE Norwegian Dairies).

REFERENCES


