175

Original article

Joint modelling of repeated measurements and event time: Application to performance traits and survival of lambs bred in sub-humid tropics

Rosemary NGUTI[a,c*], Tomasz BURZYKOWSKI[b], John ROWLANDS[c], Didier RENARD[b], Paul JANSSEN[b]

[a] Department of Mathematics, University of Nairobi, Kenya
[b] Center for Statistics, Limburgs Universitair Centrum, Diepenbeek, Belgium
[c] International Livestock Research Institute, Nairobi, Kenya

Abstract – We considered the analysis of a study for Dorper, Red Maasai and crossbred lambs born over a period of 6 years at the Diani Estate, Kenya. The study was designed to compare survival and performance traits of genotypes with differing susceptibilities to helminthiasis. The available data include information on time to death and repeated measurements of body weight, packed cell volume (PCV) and faecal egg count (FEC) of the animals. In the paper, we consider joint modelling of the survival time and the repeated measurements. Such an approach allows to account for the possible association between the survival and repeated measurement processes. The advantages and limitations of the joint modelling are discussed and illustrated using the Diani Estate study data.

repeated measurements / time-to-event / joint modelling / sheep / helminthiasis

1. INTRODUCTION

The data used in this study came from an animal breeding experiment carried out by the International Livestock Research Institute (ILRI) from 1991 to 1996 [1] at the Diani Estate, Kenya Coast. The objective of the experiment was to study genetic resistance to naturally acquired gastro-intestinal nematodes in different breeds of sheep, namely the Red Maasai (found in East Africa

---

* Corresponding author: r.wangeci@cgiar.org

and perceived to be resistant to helminthiasis), Dorper (originating from South Africa and presumed to be susceptible) and their cross breeds. Three ewe genotypes (Dorper, Red Maasai, and Red Maasai × Dorper) were mated to Dorper and Red Maasai rams in single-sire mating groups in a diallel design to generate two pure-bred and four cross-bred genotypes. The data collected during the study consisted of repeated measurements of the traits: body weight (BWT), packed cell volume (PCV) and faecal egg count (FEC), which were measured periodically over a lamb's first year of life. Overall, 1745 lambs were born alive during the six-year period of the study. The repeated measurements, however, were highly unbalanced since 655 (38%) of the lambs died and 92 (5%) were stolen before they reached one year of age.

An assessment of the level of genetic resistance has been carried out by Baker *et al.* [1, 2], by applying linear mixed models to individual measurements of the traits that were recorded from birth to one year of age. In the analysis undertaken by these authors, only the information on the animals that survived to each of the analysed time points was utilised. Nguti *et al.* [16] used these same traits in a frailty model for survival, where a correlation was induced among the ages at the time of death for lambs from the same sire. In that analysis, PCV and FEC were separately included in the model as time-varying covariates, with either a baseline or time-varying BWT. One handicap of this latter analysis was that the measured traits were only recorded at monthly time points, and were thus unknown at the time of death. To overcome this problem, the nearest preceding measured value of the trait was used to impute for the value at the time of death, resulting in a piecewise covariate process. In addition, the analysis did not correct for the possible measurement error in the recorded values of BWT, PCV and FEC. Such an error can result from both the analytical error and short-term, biological variability. Failure to account for the measurement error and for any missing time-varying covariate observations has been shown to cause the estimated regression parameters in the Cox proportional hazard (PH) model to be biased towards the null [4, 18].

In the study described, as in many longitudinal studies where individuals are followed over time, the data can be grouped into three categories: (1) the elapsed time to an event such as death; (2) repeated measurements of time-varying variables (like PCV, FEC, BWT); (3) time-varying (*e.g.*, rainfall) or constant baseline (*e.g.*, sire, breed) covariates that may affect both the repeated measurement and the time-to-event processes. When modelling of the repeated measurements is of interest, one may focus, for instance, on how the measurements change with time, on how the parameter estimates are influenced by the drop-out of individuals during the course of the study, or on how

the measurements may be affected by the additional covariates. Looking at things from a time-to-event process point of view, the interest may focus on, *e.g.*, how the time to the event is affected by both the repeated measurement process and the additional covariates. A vast amount of literature exists on the methods suitable for either approach. To analyse event-time data, parametric or non-parametric models can be used, but the Cox PH model is often the method of choice. Repeated measurements are commonly analysed using linear mixed effects models [13]. These models are attractive for several reasons, one of them being the ability to easily accommodate unbalanced designs, especially regarding the timing and frequency of the observations. The models also allow for an explicit partitioning of variability and estimation of individual effects. In particular, at least two sources of variability are readily identified: between- and within-individual variation. The between-individual variability is often modelled by a vector of correlated, individual random effects.

In the last ten years, many methods, which simultaneously use the information available in both the time-to-event and the repeated measurement processes, have been proposed in medical research. In particular, several models have been developed in the area of acquired immunodeficiency syndrome (AIDS) research [5, 7, 20, 22, 24] and in schizophrenia studies [9, 25]. A detailed review of research work in joint modelling of times to an event and repeated measurements is given in the reference [23]. The need for joint models to model survival and performance traits in animal studies is discussed in [6].

Several advantages of joint modelling of the repeated measurement and the time-to-event processes have been highlighted in the literature: (1) the repeated measurements can be extrapolated from the observed measurement times to the specific event time in a way that utilises the entire measurement history; (2) the time-to-event is allowed to depend on the 'true' but unknown value of the repeated measurement, thus making an adjustment for the measurement error, which in turn leads to reduced bias of the parameter estimates of the Cox model; and (3) the repeated measurement process is adjusted for any loss of information arising from death or loss of individuals.

The objective of the current paper was to use the joint modelling approach to model the time to death and the repeated measurements of PCV, BWT and FEC. For this, the methodology proposed by Henderson *et al.* [9] was used. The paper is organised as follows. In Section 2 more details on the motivating dataset are given. Section 3 provides a brief background on the linear mixed effects and the Cox PH models, as well as the joint model of Henderson *et al.* [9]. In Section 4 we adopt the joint model to the analysis of the motivating dataset. The results are presented in Section 5 and discussed in Section 6.

2. MOTIVATING DATA

Measurements of PCV, FEC and BWT were taken from lambs from birth to one year of age in batches of lambs born in each of the years 1991 to 1996. Packed cell volume and FEC were measured according to the methods reported by Baker *et al.* [1]. All lambs were weighed at birth and their BWT, PCV and FEC were subsequently recorded at one and two months of age. On either of these latter occasions, when individual lambs had an FEC greater than or equal to 2000 eggs per gram (epg) and/or a PCV less than or equal to 20%, they were treated (drenched) with an anthelmintic drug. Lambs with low PCV were also checked for trypanosome infections. At about three months of age, when weaned, the lambs were again weighed, and blood and faecal samples were collected for PCV and FEC, respectively. All lambs were then drenched. The lambs were then left to graze on pasture, separately from the ewes and rams. Every week, a monitor group of about 50 lambs, made up of approximately equal numbers of lambs of each genotype and sex, was sampled and their mean FEC was recorded. If the mean FEC was over 2000 epg then, during two consecutive days, all lambs were weighed, faeces and blood samples were taken and the lambs were drenched. This procedure was followed until the lambs reached on average one year of age. It resulted in five drenchings per year except for 1994 and 1996. In 1994, the year with the highest rainfall, the lambs were drenched eight times post-weaning, while in 1996 six drenchings occurred.

3. METHODOLOGY

In this section, we first briefly discuss the methods used in (separate) modelling of repeated measurements and time-to-event data. Then we shortly review joint modelling.

3.1. Linear mixed effect models for repeated measures

Data sets resulting from follow-up studies are often highly unbalanced, with subjects having an unequal number of measurements. Moreover, the data have complex correlation structure due to repeated measurements for each individual. As a result, such data are not ideally suited to analysis by classical least squares techniques and linear mixed effects models [13] are now standard tools for analysing such complex hierarchical data.

Let Y_{ij} denote the observed jth measurement for the ith individual recorded at time t_{ij} ($i = 1, \ldots, N$; $j = 1, \ldots, n_i$) and let $Y_i^T = (Y_{i1}, Y_{i2}, \ldots, Y_{in_i})$. Then a linear mixed effects model is written as

$$Y_i = X_{1i}\beta_L + Z_i s_i + \varepsilon_i, \tag{1}$$

where X_{1i} and Z_i are $n_i \times p$ and $n_i \times q$ design matrices, respectively, β_L is a $p \times 1$ vector containing the fixed effects, and s_i is a $q \times 1$ vector of the random effects. It was assumed that s_i is $N(\mathbf{0}, G)$, *i.e.*, it is normally distributed with mean zero and variance-covariance matrix $G = [g_{kl}]$, where $g_{kl} = \text{Cov}(s_{ik}, s_{il})$. Furthermore, it was assumed that s_i is independent from the vector of residual random errors ε_i. The residual errors were assumed to be $N(\mathbf{0}, R_i)$, with variance-covariance matrix R_i depending on i only *via* its size ($n_i \times n_i$). It then follows that, marginally, Y_i is normally distributed with mean $X_{1i}\beta_L$ and variance-covariance matrix $V_i = Z_i G Z_i^T + R_i$. In model (1), R_i captures the within-individual variability while the between-individual variability is modelled through the random effects s_i. In particular, if

$$Z_i^T = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ t_{i1} & t_{i2} & \ldots & t_{in_i} \end{pmatrix}$$

then model (1) is known as a random *intercept* and *slope* model (see reference [21], p. 25). The underlying assumption of this model is that the measurements increase linearly with time, but for each individual the linear trend has its own intercept and slope. Furthermore, if $\text{Var}(\varepsilon_{ij}) = \sigma_e^2$, the result that the assumed covariance function of the response for this model is

$$\text{Cov}(Y_{ik}, Y_{il}) = g_{11} + g_{22} t_{ik} t_{il} + g_{12}(t_{ik} + t_{il}) + \mathbf{1}(k = l)\sigma_e^2, \tag{2}$$

where $\mathbf{1}(A)$ is the indicator function of event A. Note that function (2) is quadratic over time.

Model (1) has been used extensively to analyse repeated measurements arising from animal breeding programmes [8, 11, 14, 15]. In these applications more emphasis has been placed on the covariance structure of the random effects (s_i, ε_i) in order to capture different sources of variability, such as those due to maternal, paternal and environmental effects.

To estimate the parameters of model (1), various approaches can be applied. One approach is the classical method of maximum likelihood (ML), which results in generalised least squares (GLS) estimates for β_L. This method of estimation, however, leads to underestimation of the variance parameters involved in G and R_i. As an alternative, the restricted maximum likelihood estimation (REML) is more commonly used, which remedies this problem [21].

3.2. The Cox proportional hazard model

Let the ith individual ($i = 1, \ldots, N$) be observed from a time zero to a failure time T_i or to a potential right censoring time C_i. Let $T_i^o = \min(T_i, C_i)$, be the observed time and δ_i be the censoring indicator which is equal to 1 if $T_i^o = T_i$ and 0 otherwise. Hence the observed data for the ith subject are (T_i^o, δ_i). The basic analytical quantities for time-to-event data are the survivor function

$$S(t) = Pr(T \geq t),$$

which is the probability of surviving beyond time t, and the hazard function

$$\lambda(t) = \lim_{\Delta t \to 0^+} \frac{Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t},$$

which is the instantaneous failure rate after surviving up to time t. In most time-to-event studies, interest focuses on how the hazard function is affected by independent covariates. To assess this effect, Cox [3] has proposed the following hazard model:

$$\lambda_i(t) = \lambda_0(t) \exp\left(x_{2i}^T \beta_S\right), \tag{3}$$

where $\lambda_0(t)$ is the baseline hazard function common to all individuals, x_{2i} is a $p \times 1$ vector of the observed covariates for the ith individual, and β_S is the corresponding vector of regression parameters to be estimated. Model (3) is known as the Cox proportional hazard model and has been used extensively over the last three decades in the analysis of failure-time data.

The Cox model assumes that the analysed time-to-event observations are independent. Models which allow to account for correlation of the failure times have been studied extensively (see [10, 12, 19]). With clustered time-to-event data, a common approach is to add a random multiplicative factor to the Cox model, shared by the observations from the same cluster. This factor is called the frailty and accounts for the correlation of the times within a cluster (*e.g.*, the correlation of survival times of lambs from the same sire). The resulting model is called a shared frailty model.

3.3. A joint model for repeated measurements and time-to-event

Let $Y_i^T = (Y_{i1}, Y_{i2}, ..., Y_{in_i})$ be the vector of the repeated measurements for the ith individual measured at times $t_i^T = (t_{i1}, \ldots, t_{in_i})$. Let $T_i^o = \min(T_i, C_i)$ and δ_i denote, respectively, the observed time and censoring indicator for

the ith individual. The observed data available for the ith individual are thus $(T_i^o, \delta_i, Y_i, t_i, X_{1i}, x_{2i})$, where X_{1i} denotes the matrix of the observed values of covariates believed to influence the repeated measurements Y_i, while x_{2i} is a vector of the observed values of covariates believed to affect the time-to-event process.

Henderson *et al.* [9] have proposed a model for the joint analysis of both the time-to-event and repeated measurements. They postulate a latent (unobserved) bivariate Gaussian process $W_i(t) = \{W_{1i}(t), W_{2i}(t)\}$ such that the repeated measurements process depends on $W_{1i}(t)$ and the event time process depends on $W_{2i}(t)$. In particular, for the repeated measurements process, consider a model of the general form

$$Y_i = \mu_i(t_i) + W_{1i}(t_i) + \epsilon_i, \tag{4}$$

where ϵ_i is an $N(0, R_i)$ error vector such that R_i is a diagonal matrix and $\text{Var}(\epsilon_{ij}) = \sigma_e^2$. Note that in (4), we use $\mu_i(t_i) = (\mu_i(t_{i1}), \ldots, \mu_i(t_{in_i}))^T$ and $W_{1i}(t_i) = (W_{1i}(t_{i1}), \ldots, W_{1i}(t_{in_i}))^T$ as shorthand notation. Furthermore, $\mu_i(t_i)$ is the systematic component, which can be described by a linear model, *e.g.*, $\mu_i(t_i) = X_{1i}\beta_L$. As a basic example for the latent process $W_{1i}(t)$, Henderson *et al.* [9] consider

$$W_{1i}(t) = U_{1i} + U_{2i}t, \tag{5}$$

where (U_{1i}, U_{2i}) is a bivariate normal random vector with zero mean and variance-covariance

$$G_1 = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

One can observe that $\mu_i(t)$ in model (4) corresponds to $X_{1i}\beta_L$ in model (1), while $W_{1i}(t)$ corresponds to $Z_i s_i$, with $s_i \equiv (U_{1i}, U_{2i})^T$.

On the contrary, the time-to-event is modelled through a Cox proportional hazard model

$$\lambda_i(t) = \lambda_0(t) \exp\left\{x_{2i}^T \beta_S + W_{2i}(t)\right\}. \tag{6}$$

It is assumed that the repeated measurement and time-to-event processes are conditionally independent given $W_i(t)$. However, in order to induce association between the two processes, $W_{2i}(t)$ is taken to be related to particular components of $W_{1i}(t)$. This is achieved *via* the general equation

$$W_{2i}(t) = \gamma_1 U_{1i} + \gamma_2 U_{2i} + \gamma_3 W_{1i}(t). \tag{7}$$

For example, a joint model with $W_{2i}(t) = \gamma_1 U_{1i} + \gamma_2 U_{2i}$, would allow both the random intercept U_{1i} and slope U_{2i}, involved in (5), to affect the risk of the event.

The parameters of the models for the repeated measurement process and the time-to-event process are estimated jointly by maximising the observed joint likelihood of the data, as described in the references [23, 24].

4. APPLICATION

We now describe the application of the joint model described in the previous section to the data introduced in Section 2. Separate analyses of the repeated measurements of PCV, FEC and BWT were performed. Survival times of lambs that survived beyond one year, or those of lambs that were stolen, were censored at one year and at the last recorded observation, respectively.

Nguti *et al.* [16] reported an average lamb mortality of 19% in the pre-weaning period and 31% in the post-weaning period. The age at death during the post-weaning period ranged from 3 to 12 months (median 6.4 months). The number of repeated measurements recorded from weaning ranged from 1 to 8 (median 6) per lamb with 1994 having the most post-weaning measurements. Variable patterns of each trait were observed over the one year as well as across the six years. For example, Figure 1 shows the scatter plot of the measurements recorded from one to 12 months for PCV. In the plot, individual profiles for a randomly selected sample of 15 lambs are highlighted. Although all animals were weighed and sampled on the same day, the ages varied as a result of the lambs being born within a period of about 20−40 days.

In the joint models with either PCV, BWT or FEC as repeated measurements, the fixed effects of genotype (6 levels), year of birth (6 levels) and sex (2 levels) were included in the repeated measurements component of the joint model. As suggested by Figure 1, PCV was assumed to be curvilinear over time. This curvilinear trend over time was also assumed for FEC and BWT (figures not shown). The age of dam (5 levels) was considered as a baseline covariate for the time-to-event component only, but not for the repeated measurements, where it was not found to be significant.

Consequently, we can define $\boldsymbol{\beta}_L^T = [\mu, \boldsymbol{\beta}_0^T]$, where $\boldsymbol{\beta}_0^T = [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4, \varsigma_5, \varpi_1]$, α_m ($m = 1, \ldots, 5$) are the binary indicators capturing the breed effects, ς_k ($k = 1, \ldots, 5$) are the indicators for year of birth effects and ϖ_l is the binary indicator for males. As a result, the repeated measurements model can be written as

$$\boldsymbol{Y}_i = \boldsymbol{X}_{1i}\boldsymbol{\beta}_L + \eta_1 \boldsymbol{t}_i + \eta_2 \boldsymbol{t}_i^* + \boldsymbol{W}_{1i}(\boldsymbol{t}_i) + \boldsymbol{\epsilon}_i, \tag{8}$$

where $\boldsymbol{X}_{1i}$ is the $n_i \times 12$ design matrix corresponding to $\boldsymbol{\beta}_L$, (η_1, η_2) are the parameters associated with the time trend, $\boldsymbol{t}_i^{*T} = (t_{i1}^2, \ldots, t_{in_i}^2)$ is the vector of the
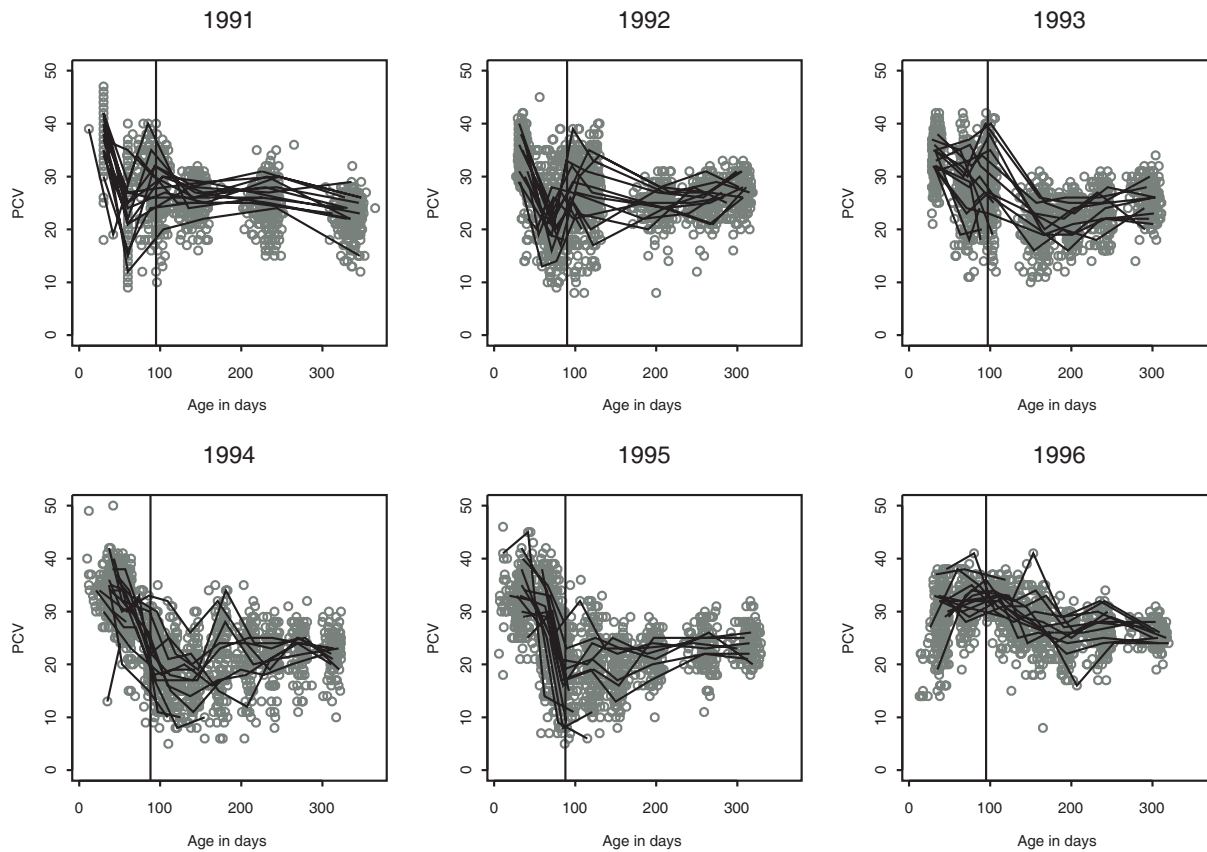
Figure 1. PCV measurements for the years 1991 to 1996. Bold vertical line indicates when the lambs were weaned.

quadratic times, and $W_{1i}(t_i) = (W_{1i}(t_{i1}), \ldots, W_{1i}(t_{in_i}))^T$. Let $\boldsymbol{\beta}^T = (\boldsymbol{\beta}_L^T, \eta_1, \eta_2)$, and $X_{1i}(t_i) = (X_{1i} \mid t_i \mid t_i^*)$ be the $n_i \times 14$ design matrix corresponding to $\boldsymbol{\beta}$. Model (8) can then be re-written as

$$Y_i = X_{1i}(t_i)\boldsymbol{\beta} + W_{1i}(t_i) + \boldsymbol{\epsilon}_i. \qquad (9)$$

To specify the survival component of the joint model, let $\boldsymbol{\beta}_S^T = [\boldsymbol{\beta}_0^T, a_1, a_2, a_3, a_4]$, where a_r $(r = 1, \ldots, 4)$ are the binary indicators coding the dam age groups (with levels: ≤ 2 years, 3, 4, 5 and ≥ 6 years). The model for survival time is then given by

$$\lambda_i(t) = \lambda_0(t) \exp\left\{x_{2i}^T \boldsymbol{\beta}_S + W_{2i}(t)\right\}, \qquad (10)$$

where x_{2i} is the $n_i \times 15$ design matrix associated with $\boldsymbol{\beta}_S$.

For all three traits the following settings for W_{1i} and W_{2i} were considered:

(S1) $W_{1i}(t) = U_{1i},$ $W_{2i} = 0;$

(S2) $W_{1i}(t) = U_{1i} + U_{2i}t,$ $W_{2i} = 0;$

(S3) $W_{1i}(t) = U_{1i},$ $W_{2i} = \gamma W_{1i};$

(S4) $W_{1i}(t) = U_{1i} + U_{2i}t,$ $W_{2i} = \gamma_1 U_{1i} + \gamma_2 U_{2i} + \gamma_3 W_{1i}(t).$

Settings (S1) and (S2) assume independence between the repeated measurement and survival processes. Settings (S3) and (S4) correspond to (S1) and (S2), respectively, with respect to the structure of $W_{1i}(t)$, but allow for dependence between the processes (joint models).

To obtain parameter estimates for the fixed effects, variance components and the association parameters of the joint models (S3) and (S4) specified above, a programme was written in SAS®. Estimates from either setting (S1) and (S2) were computed using PROC MIXED (for repeated measurements) and PROC PHREG (for survival time) in SAS®.

Estimates of the standard errors for all parameter estimates in the joint models were obtained by using the jackknife method. This was achieved by leaving out the observations for lambs from the same sire and then re-fitting the model to the remaining observations. Classically, jackknife provides reliable estimates of the standard errors if the observations omitted are independent from those that are left in. When observations for lambs from the same sire are left out, so is the genetic component of these lambs. This genetic component is assumed here to provide more individual contribution to the lamb characteristics (*e.g.* survival, BWT) than the common environmental components, which are shared by lambs born in the same year. This is supported by the findings in reference [2]. These authors show that for the analysed data set, the differences observed in heritability estimates of PCV and FEC for Dorper- compared with Red Maasai-sired lambs were more likely due to the differences in genetic variance rather than in environmental variability.

5. RESULTS

In this section the results of the fitted models for PCV, BWT and FEC for settings (S1)–(S2) are reported. For each trait the results of the joint model are compared with those of the corresponding independence model for both the repeated measurements and survival estimates.

5.1. Packed cell volume from one month

Initially, the analysis of the repeated measurements for PCV from one month until one year of age was considered. When fitting the model corresponding to setting (S2), a non-positive definite estimate of the variance-covariance matrix G_1 (see Sect. 3.3) was obtained. On further investigation, it was discovered that the PCV repeated measurements were negatively correlated, with the correlation increasing in absolute value over time. This negative serial correlation cannot be captured by a model with a random intercept and random slope, as specified under setting (S2). In particular, by taking into account the magnitude of the correlation and low variability of the slopes (g_{22} in Eq. (2)) of individual profiles, the variance of the random slope is estimated to be less than or equal to 0, which is obviously an inadmissible value. Therefore, for the PCV measurements collected from one month to one year of age, the models for settings (S2) and (S4) could not be fitted. On the contrary, the models for settings (S1) and (S3), which do not account for the negative correlation, could be used. The results for these models are given in Table I. Strictly speaking, one should treat the results with caution since they are based on models with a possibly misspecified variance-covariance structure.

Setting (S1): Under this setting, which assumes independence between PCV measurements and survival time, the Dorper (D×D) breed had the lowest mean PCV from one month to one year of age, which was between 0.1 to 1.9% units lower than for other genotypes (see Tab. I). This difference increased as the Red Maasai genotype in the lambs increased, with the Red Maasai having the highest mean PCV. The linear and quadratic time effects were both significant ($P < 0.001$) implying an average non-linear trend in PCV. The trend is as indicated in Figure 1, which shows a general sharp decline in PCV after one month in all years except 1996 followed by a slight rise. The lambs born in 1992–1995 had on average lower PCV (0.3 to 3.2%) units than those born in 1991. The mean PCV was the highest in 1996. On average, male lambs had lower PCV than female lambs.

By exponentiating the estimates given in the 'Survival model - S1' column in Table I, one can see that, as compared to the Dorper, the relative

Table I. Estimates from settings S1 ($W_{1i}(t) = U_{1i}$, $W_{2i}(t) = 0$) and S3 ($W_{1i}(t) = U_{1i}$, $W_{2i}(t) = \gamma W_{1i}(t)$) for repeated measurements of PCV(%) from one month to 12 months and survival.

	Repeated measurements model				Survival model			
	S1		S3		S1		S3	
	est	s.e.	est	s.e.	est	s.e.	est	s.e.
Fixed effects								
Intercept	31.730	0.238	31.457	0.394	-	-	-	-
time(months)	−2.856	0.052	−2.866	0.119	-	-	-	-
time*time	0.217	0.005	0.217	0.011	-	-	-	-
Genotype								
DxD	ref		ref		ref		ref	
Dx(DxR)	0.132	0.231	0.237	0.281	−0.476	0.121	−0.514	0.136
DxR	0.355	0.332	0.449	0.392	−0.569	0.172	−0.583	0.209
RxD	1.406	0.260	1.603	0.381	−0.872	0.169	−0.945	0.197
Rx(RxD)	1.668	0.224	1.935	0.313	−1.341	0.141	−1.390	0.184
RxR	1.937	0.279	2.230	0.356	−1.342	0.177	−1.407	0.226
Year of birth								
1991	ref		ref		ref		ref	
1992	−0.412	0.222	−0.373	0.303	−0.023	0.185	−0.048	0.209
1993	−0.340	0.219	−0.470	0.248	0.921	0.149	0.948	0.157
1994	−2.867	0.252	−3.098	0.270	1.358	0.167	1.361	0.185
1995	−3.178	0.251	−3.384	0.291	1.328	0.165	1.361	0.187
1996	1.097	0.257	1.008	0.306	0.724	0.197	0.769	0.248
Gender								
Females	ref		ref		ref		ref	
Males	−0.512	0.138	−0.558	0.156	0.220	0.088	0.232	0.115
Age of dam								
≤2 years	-	-	-	-	ref		ref	
=3 years	-	-	-	-	−0.277	0.158	−0.271	0.160
=4 years	-	-	-	-	−0.544	0.167	−0.505	0.181
=5 years	-	-	-	-	−0.368	0.159	−0.278	0.182
≥6 years	-	-	-	-	−0.509	0.168	−0.461	0.150
Variances								
σ_e^2	25.817	0.398	25.905	0.839	-	-	-	-
σ_1^2	3.037	0.283	2.804	0.327	-	-	-	-
Association								
γ	-	-	-	-	-	-	−0.236	0.020

σ_e^2 is residual error variance; σ_1^2 is the only variance component in matrix $\boldsymbol{G}_1$.

mortality hazard of the other genotypes ranged from $\exp(-0.476) = 0.62$ to $\exp(-1.34) = 0.26$. The R × (R × D) and R × R breeds had the lowest, and similar, mortality. The hazard of the lambs born in the years 1993–1996 was statistically significantly higher than that of the lambs born in 1991 and ranged from 2.2 to 4.0. Male lambs had a higher mortality hazard than females while the hazard ratio decreased with increasing age of the dam.

Setting (S3): This setting corresponds to (S1), but assumes dependence between PCV measurements and survival time. As compared with (S1) the differences in the mean PCV, relative to the Dorper breed, increased slightly for all other genotypes. For instance, the estimated mean PCV for the non-Dorper genotypes was 0.2−2.2% units higher than for the Dorper breed (Tab. I). This increase might be the result of the adjustment of the analysis of the repeated measurements for the variation in death rates. The estimated time trend parameters for the repeated measurements model for the (S1) setting were similar to those obtained for (S3).

Relative to the mortality hazard for the Dorper breed, the hazard ratio for the non-Dorper genotypes now ranged between 0.60 to 0.24, as compared to the (S1) setting (Tab. I). Significant negative estimates ($P < 0.001$) were obtained for the association parameters (γ in Tab. I) for the survival model under (S3). This indicates that the mortality hazard decreased with increasing PCV.

5.2. Packed cell volume from weaning

Since the critical period for assessing genetic resistance to endoparasites in lambs is between weaning and 12 months of age [1], the analysis of the PCV repeated measurements for the period from weaning onwards was also considered. In this analysis, the survival time was redefined by using weaning as the time of origin. Consequently in this analysis only the animals alive at the time of weaning were considered. Models under settings (S1) to (S4) were considered. The results for settings (S1) and (S3) (not shown) were similar to those reported for the analysis of data from one month of age (Tab. I). However, unlike the period from one month of age, the time trend had a more moderate negative slope estimate. This corresponds to Figure 1, which shows a more gradual decline in PCV after weaning than before. Furthermore, the relative mortality hazard in the post-weaning period exhibited similar patterns as in the analysis of data from one month of age, but now the R × (R × D) had the lowest mortality ($\exp(-1.641) = 0.19$) when compared to the Dorper breed. A detailed description of the results from (S2) and (S4) (Tab. II) is given below.

Settings (S2): In the repeated measurements model with both random intercepts and slopes, (Tab. II), similar trends for the fixed effects parameter estimates were observed as in the simpler, random-intercept-only model. However the ranges of the estimates were reduced. For example for setting (S1), the genotype parameter ± s.e. estimates were the following: 0.361 ± 0.287 (D × (D × R)), 0.978 ± 0.412 (D × R), 1.756 ± 0.318 (R × D), 2.405 ± 0.276 (R × (R × D)) and 2.866 ± 0.344 (R × R). On the contrary, for setting (S2) the corresponding estimates were (see Tab. II): 0.169 ± 0.267 (D × (D × R)), 0.726 ± 0.387 (D × R), 1.585 ± 0.290 (R × D), 2.070 ± 0.254 (R × (R × D)) and 2.586 ± 0.319 (R × R).

The random intercept and slope were negatively correlated ($\sigma_{12} = -1.82$, see Tab. II). This implies that lambs with a high PCV at weaning had a more rapid decline in PCV than those with a low PCV. The estimated variance component for the random intercept (σ_1^2) was 15.89 (s.e. 0.99) compared to 6.47 (s.e. 0.43) in the simpler model (S1) (results not shown). Note, however, that, according to equation (2), the negative correlation between random intercept and slope will reduce the total variability under the (S2) setting, making it similar to the sum of σ_1^2 and σ_e^2 in the simpler (S1) model. It is also worth noting that, according to equation (2), under (S2) the covariance between two PCV measurements depends on time, and can be negative if the correlation between random intercept and slope is negative. Since it is possible that the correlation between PCV measurements may change with time and may be negative (see remarks in Sect. 5.1), this gives more credibility to the results obtained for (S2) because the model used in an (S1) settting fimposes a constant positive correlation between PCV measurements.

Setting (S4): As compared with (S1) and (S2) settings, the differences in the mean PCV for all other genotypes, as compared with the Dorper breed, increased slightly in the joint models constructed under both (S3) (results not shown) and (S4) (Tab. II) settings. For instance, for the (S3) setting, the estimated mean PCV from weaning for the non-Dorper genotypes was 0.6−3.4% units higher than for the Dorper breed. For the (S4) setting (Tab. II), the difference was between 0.3% and 3.0%. This increase might be attributed to the adjustment of the analysis of the repeated measurements for the variation in death rates.

Relative to the mortality hazard for the Dorper breed, the hazard ratio for the non-Dorper genotypes now ranged from 0.55 to 0.16 for the (S4) setting (Tab. II). These estimates were lower than the estimates obtained for the corresponding (S1) and (S2) survival models. For (S3) setting the mortality hazard for lambs born in 1993–1995 was about five times higher than in 1991

Table II. Estimates from settings S2 ($W_{1i}(t) = U_{1i} + U_{2i}t$, $W_{2i}(t) = 0$) and S4 ($W_{1i}(t) = U_{1i} + U_{2i}t$, $W_{2i}(t) = \gamma_1 U_{1i} + \gamma_2 U_{2i} + \gamma_3 W_{1i}(t)$) for repeated measurements of PCV(%) from weaning to 12 months and survival.

| | Repeated measurements model | | | | Survival model | | | |
| | S2 | | S4 | | S2 | | S4 | |
	est	s.e.	est	s.e.	est	s.e.	est	s.e.
Fixed effects								
Intercept	25.817	0.273	25.844	0.444	-	-	-	-
time (months)	−1.465	0.052	−1.562	0.081	-	-	-	-
time*time	0.160	0.006	0.164	0.009	-	-	-	-
Genotype								
DxD	ref	-	ref	-	ref	-	ref	-
Dx(DxR)	0.169	0.267	0.333	0.354	−0.492	0.142	−0.594	0.200
DxR	0.726	0.387	0.842	0.489	−0.678	0.200	−0.634	0.284
RxD	1.585	0.290	1.892	0.451	−0.975	0.195	−1.221	0.277
Rx(RxD)	2.070	0.254	2.521	0.404	−1.641	0.176	−1.843	0.257
RxR	2.586	0.319	2.991	0.440	−1.332	0.201	−1.569	0.304
Year of birth								
1991	ref	-	ref	-	ref	-	ref	-
1992	1.460	0.237	1.392	0.404	0.037	0.228	−0.260	0.340
1993	−0.779	0.239	−1.208	0.358	1.468	0.174	1.678	0.282
1994	−3.685	0.286	−4.311	0.631	1.570	0.203	1.343	0.403
1995	−3.631	0.290	−4.285	0.661	1.533	0.203	1.396	0.471
1996	2.101	0.286	1.998	0.365	0.531	0.268	0.582	0.308
Gender								
Females	ref	-	ref	-	ref	-	ref	-
Males	−0.539	0.154	−0.597	0.188	0.300	0.104	0.357	0.170
Age of dam								
≤2 years	-	-	-	-	ref	-	ref	-
=3 years	-	-	-	-	−0.377	0.177	−0.425	0.184
=4 years	-	-	-	-	−0.728	0.187	−0.685	0.205
=5 years	-	-	-	-	−0.549	0.178	−0.453	0.227
≥6 years	-	-	-	-	−0.810	0.196	−0.676	0.213
Variances								
σ_e^2	12.901	0.272	12.641	0.401				
σ_1^2	15.894	0.990	15.380	1.517				
σ_{12}	−1.817	0.145	−1.590	0.236				
σ_2^2	0.276	0.024	0.261	0.028				
Association								
γ_1	-	-	-	-	-	-	−0.273	0.097
γ_2	-	-	-	-	-	-	−1.986	0.505
γ_3	-	-	-	-	-	-	−0.251	0.087

σ_e^2 is residual error variance; σ_1^2, σ_{12} and σ_2^2 are variance-covariance elements of matrix $\boldsymbol{G}_1$.

(with parameter ± s.e. estimates as: 1.637±0.209 (1993), 1.640±0.252 (1994), 1.677±0.250 (1995)). For (S4) the ratio was about five for 1993, while for 1994 and 1995 it was about four times, a lower ratio than for (S3). This is supported by Figure 1. The lambs born in 1994 and 1995 had much lower PCV measurements at weaning than those born in 1991, and PCV has been shown to correlate with survival. However for the former, PCV increased over time while those for the latter decreased. Thus, the adjustment for the PCV evolution over time results in a slight decrease in the mortality hazard for 1994–1995. On the contrary, the lambs born in 1993 and 1991 had almost similar PCV measurements at weaning. The decrease in 1993 over time was, however, much sharper (larger negative slope) than in 1991. Adjusting for this decrease translates into a higher mortality hazard for 1993. Finally, higher hazard ratios are observed for 1994 and 1995 in (S3) than (S4), since the latter model adjusts the risk only for the overall level of PCV over time (without adjusting for the rate of change).

Significant negative estimates ($P < 0.001$) were obtained for all the association parameters: $\hat{\gamma} = -0.303 \pm 0.0203$ for setting (S3), and $\gamma_1 - \gamma_3$ as in Table II for setting (S4). This indicates that the mortality hazard decreased with increasing PCV. Thus, after weaning in the (S3) setting, lambs with PCV measurements higher than the average had a lower mortality hazard than those with lower PCV measurements. The standard deviation of the distribution of the random intercepts in the repeated measurements part of the joint model for the (S3) setting was estimated to be equal to 2.51. Thus, the model predicts that for a (random) increase of PCV due to an increase of the random intercept $W_{1i} = U_{1i}$ by one standard deviation, the risk of death decreases by $\exp(-\hat{\gamma} * W_{1i}) = \exp(-0.303 * 2.51) = 0.47$. For the (S4) model, a large negative estimate was obtained for $\gamma_2(= -1.92)$, which corresponds to the random individual slope. Thus, the model indicates that at any time lambs that had a large decrease in PCV had an increased risk of death.

5.3. Body weight

Measurements of body weight from birth to one year of age were used in this analysis. The models were fitted using all four settings. The parameter estimates of the fixed effects were similar for settings (S1)–(S3) and (S2)–(S4). We thus report only the estimates of the models constructed under settings (S2) and (S4) which are shown in Table III.

Table III. Estimates from settings S2 ($W_{1i}(t) = U_{1i}+U_{2i}t$, $W_{2i}(t) = 0$) and S4 ($W_{1i}(t) = U_{1i} + U_{2i}t$, $W_{2i}(t) = \gamma_1 U_{1i} + \gamma_2 U_{2i} + \gamma_3 W_{1i}(t)$) for repeated measurements of BWT(kg) from birth to 12 months and survival.

| | Repeated measurements model | | | | Survival model | | | |
| | S2 | | S4 | | S2 | | S4 | |
	est	s.e.	est	s.e.	est	s.e.	est	s.e.
Fixed effects								
Intercept	4.662	0.103	4.665	0.146	-	-	-	-
time (months)	2.330	0.016	2.229	0.045	-	-	-	-
time*time	–0.101	0.001	–0.101	0.003	-	-	-	-
Gentotype								
DxD	ref	-	ref	-	ref	-	ref	-
Dx(DxR)	–0.036	0.103	–0.042	0.109	–0.494	0.108	–0.498	0.130
DxR	–0.405	0.150	–0.396	0.159	–0.630	0.157	–0.577	0.165
RxD	–0.058	0.117	–0.074	0.131	–0.809	0.149	–0.947	0.194
Rx(RxD)	–0.361	0.100	–0.391	0.125	–1.263	0.124	–1.221	0.160
RxR	–0.715	0.127	–0.727	0.142	–1.302	0.157	–1.404	0.203
Year of birth								
1991	ref	-	ref	-	ref	-	ref	-
1992	–0.621	0.102	–0.637	0.160	0.073	0.161	0.158	0.163
1993	0.504	0.100	0.647	0.138	0.763	0.136	0.017	0.219
1994	–1.852	0.120	–1.836	0.140	1.384	0.148	1.330	0.162
1995	–1.238	0.120	–1.167	0.145	1.246	0.149	0.968	0.188
1996	–2.033	0.118	–2.088	0.060	0.763	0.173	1.084	0.325
Gender								
Females	ref	-	ref	-	ref	-	ref	-
Males	0.113	0.064	0.095	0.118	0.202	0.079	0.387	0.097
Age of dam								
≤2 years	-	-	-	-	ref	-	ref	-
=3 years	-	-	-	-	–0.187	0.144	0.096	0.147
=4 years	-	-	-	-	–0.490	0.152	0.086	0.162
=5 years	-	-	-	-	–0.333	0.145	0.291	0.163
≥6 years	-	-	-	-	–0.433	0.152	0.098	0.148
Variances								
σ_e^2	2.533	0.038	2.520	0.091	-	-	-	-
σ_1^2	0.825	0.058	0.867	0.052	-	-	-	-
σ_{12}	0.179	0.013	0.221	0.017	-	-	-	-
σ_2^2	0.070	0.004	0.082	0.007	-	-	-	-
Association								
γ_1	-	-	-	-	-	-	–0.086	0.389
γ_2	-	-	-	-	-	-	–3.285	1.609
γ_3	-	-	-	-	-	-	0.006	0.079

σ_e^2 is residual error variance; σ_1^2, σ_{12} and σ_2^2 are variance-covariance elements of matrix $\boldsymbol{G}_1$.

Setting (S2): The Dorper (D × D) breed had the highest mean BWT, which was between 0.04 and 0.72 kg higher than for the other genotypes (Tab. III). There was a non-linear trend in the change of body weight over time. On average, the lambs born in 1994–1996 were lighter than those born in 1991. Male lambs were on average 0.11 kg heavier than the females.

Compared with the Dorper, the relative mortality hazard of the other genotypes ranged from 0.61 to 0.27, with the Red Maasai (R × R) and R × (R × D) having the lowest hazard. As for the previous PCV analyses, an increased mortality hazard was noted for the years 1993–1996. Lambs born to ewes ≥ 3 years of age had lower hazard than those born to younger ewes.

Setting (S4): Adjusting the repeated measurement process for the variation in death rates only had a slight effect on the parameter estimates of the (S4) models when compared to setting (S2).

The relative mortality of the genotypes now ranged from 0.61 to 0.24 with the R×R genotype having the lowest hazard mortality. This result indicates that despite being lighter in body weight when compared to the other genotypes, the Red Maasai demonstrated better performance in terms of survival. The age of the dam effect was non-significant. This could be due to the fact that in this analysis we accounted for the low body weight of lambs born to young dams, which biologically may be due to low milk production of the dam in her first parity.

Negative estimates of the parameters relating the random components of the repeated measurements model to the survival model were observed (see the estimates for $\gamma_1 - \gamma_3$ in Tab. III). In particular, there was a significant negative association only between random growth rate ($\gamma_2 (= -3.28)$, $P < 0.001$) and risk of death. This shows that the animals that had weight profiles with increasing slope had a reduced risk of death. In a reduced model with γ_1 and γ_3 constrained to zero (results not shown), the estimate obtained for the association parameter γ_2 was $\hat{\gamma}_2 = -3.262$ (s.e. $= 0.389$). The standard deviation of the random slope in this reduced model was equal to 0.22. Thus for every increase by one standard deviation, the mortality hazard associated with the rate of change in BWT was reduced by $\exp(-3.262 \times 0.22) = 0.49$.

5.4. Faecal egg count

The repeated measurements for FEC were log-transformed into LFEC = ln(FEC + 25) since the data were highly skewed. Initially, an analysis of the repeated measurements from one month to one year of age was considered. As in the case with PCV, a negative correlation between repeated measurements, that

increased over time, was observed. This problem precluded the estimation of any of the settings (S1)–(S4). Consideration of only the LFEC measurements from weaning onwards did not resolve the problem, since the negative correlation continued to persist among measurements collected towards the end of the one year period. The problem may be due to strong oscillations of the individual patterns that were observed in the LFEC measurements from one month onward, as well as from the time of weaning, that could be linked with the treatment protocol instigated in the study. A possible solution might be to extend the random structure (W_{1i}) of the repeated measurements model by including a serial correlation component in the model for the repeated measurements process. Owing to the limitation of software availability, this modelling option was not attempted.

6. DISCUSSION

The data discussed in this paper were previously analysed by Baker *et al.* [2] and Nguti *et al.* [16]. Baker *et al.* [2] analysed the repeated measurements of BWT, PCV and FEC without taking into account the survival pattern of the animals, and chose to analyse the data for each time point separately. They confirmed the higher resistance (lower FEC) and higher resilience (higher PCV) of Red Maasai than the Dorpers. Nguti *et al.* [16], studied the survival of each genotype and introduced the effects of BWT, PCV and FEC as time-varying covariates in a shared frailty model, with the frailty defined as a random effect of sire. Introduction of PCV and FEC as time-varying covariates in that analysis in models with BWT (time-invariant or time-varying) reduced the magnitude of the sire variance, confirming the moderate levels of heritability reported by Baker *et al.* [2].

In the analysis presented in the current paper, individual repeated measurements were analysed jointly with the survival process. By doing so, parameter estimates in both components of the joint model in general increased in absolute order of magnitude, as compared with the models assuming independence between the two processes. For instance, in the joint model with PCV as the repeated measurement, the range of the mortality hazard ratios for different genotypes relative to the Dorper changed from 0.61−0.19 for (S1) and (S2) models to 0.55−0.16 for (S3) and (S4) models. Thus, the adjustment for the evolution of PCV resulted in a clearer separation between the Dorpers and the Red Maasai. On the contrary, adjustment for the death rates widened the ranges of the parameters reflecting the differences among the various genotypes in PCV measurements. As shown by Baker *et al.* [2] the PCV values among the genotypes were similar at the first sampling at one month, prior to any sign

of the disease. The subsequent reductions in PCV were inversely correlated with the levels of FEC, suggesting that the differences in overall mean PCV values among the genotypes were more likely the result of the different levels of resistance between the Dorpers and the Red Maasai, rather than any genetic differences between breeds in their normal PCV values.

In general, repeated measurements such as PCV, BWT or FEC are only recorded at specific time points. When such variables are used in a proportional hazards model as time-varying covariates, the standard method is to impute the missing observations by using the last observed value. This results in a piece-wise constant profile. This approach has been shown to lead to biased model parameters [18], with the presence of measurement error in the covariate attenuating the estimates towards zero. On the contrary, in the joint analysis, the repeated measurements are imputed by the 'true' values predicted under the model for the repeated measurement process. This reduces the attenuation and can explain the increase in the magnitude of parameter estimates observed for the joint models. In addition, the joint modelling allows other characteristics of the repeated measurements pattern, such as the rate of change (slope), to influence the risk of death. The effects of these characteristics is well demonstrated in this study when repeated measurements of PCV recorded from weaning were considered. Including a random slope had a major effect on the post-weaning risks of death for the years 1993 to 1995 when compared with 1991.

Nguti [17] compares the results from the survival component of the joint model under setting (S4) for the repeated measurements of PCV (from weaning) and BWT (from birth) together with those for a Cox PH model with a time-varying covariate for these measurements. Differences in hazard estimates were observed by the two approaches. However, it would be erroneous to make direct comparisons of the two sets of results since the methodological aspects of the two models are different. The survival component of the joint model can be viewed as a conditional model (conditioned on the random effects which may be partially confounded with genotype), while a Cox PH model with a time-dependent covariate can be viewed as a population average model.

The aforementioned potential of the joint models for a more accurate analysis of the repeated measurements and survival processes is an attractive feature of the methodology. However, it might be that the models do not have the capacity to fit the actual data as well as with the FEC measurements. In such a case a careful reflection on the source of this lack-of-fit should result in proposing a change in the structure of the model. For instance, it should be noted that the post-weaning measurements were obtained at the time lambs

were diagnosed for treatment based on the observations of the sentinel group of 50 lambs (see Sect. 2). Treatment will have resulted in decreases in FEC especially, and also increases in PCV. Thus data collected in between treatments, when measurements were more likely to have been within normal limits, were missing. These regular treatments are likely to have influenced the covariance structure resulting in the observed negative correlations. A similar problem occurred pre-weaning when individual animals were treated on the basis of PCV and FEC. It should also be noted here that not all lambs that died did so due to helminthiasis. Although Nguti [16] found a difference in survival rate between Dorper and Red Maasai regardless of whether the deaths due to causes other than helminthiasis were omitted or not, the decreases in PCV and increases in FEC, that are typical of the onset of helminthiasis, may not have been reflected in deaths due to other causes. More frequent sampling and use of the data collected mid-way between treatment interventions may have both helped in better understanding the patterns of changes in PCV and FEC, thus overcoming the problem of a negative correlation.

In the application of any statistical method one needs to be aware of the assumptions made. For example, in the joint models considered in this paper, the models used a bivariate Gaussian process to induce the association between the repeated measurements and the survival process. This, in turn, imposes a particular, hierarchical structure (*e.g.*, random intercepts and slopes) of the model for the repeated measurements process. Such a structure may not necessarily be adequate for a particular dataset. As already noted, the models with random intercepts and random slopes were not able to capture the intrinsic patterns of FEC induced by the treatment interventions. On the contrary, they coped better with the BWT profiles from the time of birth. In general, random intercepts and slopes provide a representation of the dominant part of the evolution of the profiles, but do not capture a more subtle behaviour (*e.g.* short-term oscillations around an average pattern). Such a behaviour might be captured by using an autocorrelated stochastic process. Henderson *et al.* [9] has proposed the use of a non-stationary Gaussian process in their approach. Unfortunately, due to the lack of appropriate software, this solution is not yet available in practice.

The joint models applied in this paper did not allow for adjusting of correlations among survival times of lambs from the same sire. For this purpose, Nguti *et al.* [16] used frailty models (see Sect. 3.2) with a random sire effect. In theory, the models formulated by Henderson *et al.* [9] allow for the inclusion of a frailty term in the time-to-event component of the joint model, but implementation needs to await the development of appropriate software.

7. CONCLUSION

Classically, the effect of repeated measurements on the risk of death is assessed by treating the repeated measurement as a time-dependent covariate in a survival model. Estimates from such an analysis have been shown to be biased towards zero, thus showing over-estimated hazard ratios. The bias can be removed by using joint models. This paper illustrates the application of such models. Accounting for the survival (drop-out) mechanism was shown to affect the parameter estimates related to the repeated measurement process.

REFERENCES

[1] Baker R.L., Mwamachi D.M., Audho J.O., Aduda E.O., Thorpe W., Genetic resistance to gastro-intestinal nematode parasites in Red Maasai, Dorper and Maasai × Dorper ewes in the sub-humid tropics, Anim. Sci. 69 (1999) 335–344.

[2] Baker R.L., Nagda S., Rodriguez-Zas S.L., Southey B.R., Audho J.O., Aduda E.O., Thorpe W., Resistance and resilience to gastro-intestinal nematode parasites and productivity of Red Maasai, Dorper and Red Maasai × Dorper crossbred lambs in the sub-humid tropics, Anim. Sci. 76 (2003) 119–136.

[3] Cox D.R., Regression models and lifetables (with discussions), J. Roy. Stat. Soc. series B 39 (1972) 1–38.

[4] Dafni U.G., Tsiatis A.A., Evaluating surrogate markers of clinical outcome when measured with error, Biometrics 54 (1998) 1445–1462.

[5] DeGrutolla V., Tu X.M., Modelling progression of CD4-lymphocytes count and its relation to survival time, Biometrics 50 (1994) 1003–1014.

[6] Ducrocq V., Topics that may deserve further attention in survival analysis applied to dairy cattle breeding: some suggestions, in: Proceedings of the International Workshop on EU Concerted Action Genetic Improvement of Functional Traits in Cattle (GIFT): Longevity, Inter. Bull. 21 (1999) 180–189.

[7] Faucett C., Thomas D., Simultaneously modelling censored data and repeatedly measured covariates, Stats. Med. 15 (1996) 1663–1685.

[8] Foulley J.L., Quaas R.L., Heterogeneous variances in Gaussian linear mixed models, Genet. Sel. Evol. 27 (1995) 211–225.

[9] Henderson R., Diggle P., Dobson A., Joint modelling of longitudinal measurements and event time data, Biostatistics 1 (2000) 465–480.

[10] Hougaard P., Analysis of multivariate survival data, Springer, 2000.

[11] Jamrozik J., Schaeffer L.R., Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins, J. Dairy Sci. 80 (1997) 762–770.

[12] Klein J.P., Moeschberger M.L., Survival Analysis: Techniques for censored and truncated data, Springer, 1997.

[13] Laird N.M., Ware J.H., Random-effects models for longitudinal data, Biometrics 38 (1982) 963–974.

[14] Meyer K., Variance components due to direct and maternal effects for growth traits of Australian beef cattle, Livest. Prod. Sci. 31 (1992) 179–204.

[15] Meyer K., Estimates of genetic and phenotypic covariance functions for post-weaning growth and mature weight of beef cows, J. Anim. Breed. Genet. 116 (1999) 181–205.

[16] Nguti R., Janssen P., Rowlands G.J., Audho J.O., Baker R.L., Survival of Red Maasai, Dorper and crossbred lambs in the sub-humid tropics, Anim. Sci. 76 (2003) 3–17.

[17] Nguti R., Random effects survival models applied to animal breeding data, Unpublished Ph.D. dissertation, Limburgs Universitair Centrum, Belgium, 2003.

[18] Prentice R.L., Covariate measurement errors and parameter estimation in a failure time regression model, Biometrika 69 (1982) 331–342.

[19] Therneau T.M., Grambsch P.M., Modelling survival data: extending the Cox model, Springer, 2000.

[20] Tsiatis A.A., DeGruttola V., Wulfsohn M.S., Modelling the relationship of survival to longitudinal data measured with error. An application to survival and CD4 counts in patients with AIDS, J. Amer. Stat. Ass. 90 (1995) 27–37.

[21] Verbeke G., Molenberghs G., Linear Mixed Models for Longitudinal Data, Springer, 2000.

[22] Wang Y., Taylor J.M.G., Jointly modelling longitudinal and event time data with application to acquired immunodeficiency syndrome, J. Amer. Stat. Ass. 96 (2001) 895–905.

[23] Wood A., Joint modelling of longitudinal and time-to-event data, Unpublished Ph.D. dissertation, Lancaster University, UK, 2002.

[24] Wulfsohn M.S., Tsiatis A.A., A joint model for survival and longitudinal data measured with error, Biometrics 53 (1997) 330–339.

[25] Xu J., Zeger S.L., Joint analysis of longitudinal data comprising of repeated measurements and time to events, Appl. Stat. 50 (2001) 375–387.