Clustered multistate models with observation level random effects, mover–stayer effects and dynamic covariates: modelling transition intensities and sojourn times in a study of psoriatic arthritis

Sean Yiu, Vernon T. Farewell and Brian D. M. Tom

University of Cambridge, UK

Summary. In psoriatic arthritis, it is important to understand the joint activity (represented by swelling and pain) and damage processes because both are related to severe physical disability. The paper aims to provide a comprehensive investigation into both processes occurring over time, in particular their relationship, by specifying a joint multistate model at the individual hand joint level, which also accounts for many of their important features. As there are multiple hand joints, such an analysis will be based on the use of clustered multistate models. Here we consider an observation level random-effects structure with dynamic covariates and allow for the possibility that a subpopulation of patients is at minimal risk of damage. Such an analysis is found to provide further understanding of the activity–damage relationship beyond that provided by previous analyses. Consideration is also given to the modelling of mean sojourn times and jump probabilities. In particular, a novel model parameterization which allows easily interpretable covariate effects to act on these quantities is proposed.

Keywords: Clustered processes; Jump probabilities; Mover–stayer model; Multistate model; Psoriatic arthritis; Sojourn times

1. Introduction

In psoriatic arthritis, manifestations of the disease typically result in joints becoming swollen and/or painful (active joints), which are reversible through treatment or management strategies or spontaneously, and may lead to permanent joint damage. The interplay between disease activity (as measured by activity in the joint) and damage is believed to be of a causal nature with a previous investigation performed by O'Keeffe *et al.* (2011) providing an extensive discussion on the topic. In that analysis, among others, individual joint level three-state models consisting of a not-active and not-damaged state, active and not-damaged state and an absorbing damaged state were proposed and produced strong evidence of a greatly increased transition rate to damage when a joint is active (compared with a joint being not active). The three-state models were fitted under a working independence assumption (the three-state processes are independent within an individual) with a robust covariance matrix used to adjust standard errors (Lee and Kim, 1998). The purpose of this paper is to extend the current modelling framework so that greater confidence with regard to the association between activity and damage can be achieved, and also to inform on other important clinical questions.

From a statistical point of view, it is important to adjust for observed and, where possible, unobserved characteristics which are believed to be strongly related to the processes of interest, i.e. confounder variables. An analysis that does not may produce spurious associations between included covariates and the outcome. Therefore, as extensions to O'Keeffe *et al.* (2011), dynamic covariates (covariates which describe important aspects of previous developments of the process) are included to allow current transition intensities to depend on previous history (relaxing the Markov assumption), random effects are introduced into the transition intensities to account for unobserved heterogeneity and provide a more efficient estimation procedure, and a mover–stayer model (Frydman, 1984) is considered to allow for the possibility that some patients (stayers) have no propensity to develop damaged joints. Whereas much research has focused on the effect of disease activity on joint damage, no research has yet considered the reverse association (i.e. the effect of damage on activity). Specifically, it is of interest to investigate whether the disease activity process changes with onset of damage, and how if so. To inform on the possible association, the absorbing damaged state is further subdivided into an active and damaged state and a not-active and damaged state, thereby allowing the disease activity process to be modelled even after a joint has become damaged. The resulting model then utilizes the entire data set, as opposed to previously where the disease activity process was stopped once a joint had become damaged. By considering a mover–stayer model, it is also possible to investigate whether the activity process is different between movers (those who have the propensity to develop damaged joints) and stayers, and this will also contribute new knowledge towards the relationship between damage and activity.

Clustered progressive multistate models constructed using random effects have previously been proposed in the panel data literature. See, for example, Cook *et al.* (2004), O'Keeffe *et al.* (2012) and Sutradhar and Cook (2008). In our context, these models introduce time invariant, possibly multivariate random effects at the patient level to account for the correlation between joints from the same patient, time invariant unobserved heterogeneity and relaxation of the Markov assumption. A novel feature of our work is the use of observation level multivariate random effects in clustered non-progressive multistate models to account for correlation and time varying unobserved heterogeneity and the introduction of dynamic covariates to relax the Markov assumption explicitly. The proposing of this random-effects structure was motivated by the extensive lengths of follow-up and the possibly non-predictable changes in unobserved heterogeneity due to treatment or management strategies employed by the clinic and the spontaneous nature of joint activity. Such observations are less likely to result in unobserved heterogeneity being time invariant or time varying but deterministic (which is enforced by patient level random effects). Along with generalized estimating equations, copulas (Diao and Cook, 2014) and expanded state space models (Tom and Farewell, 2011) have also been proposed to handle clustering. Although there are considerable advantages to such models, they are particularly difficult to formulate and implement when more than two intermittently observed non-progressive multistate processes are of interest.

The natural multistate modelling parameterization allows covariates to act on transition intensities in a proportional hazards framework. Therefore easily interpretable covariate effects on these transition intensities can be obtained. Another natural way to view a multistate process is in terms of its sojourn times (the time spent in a state before a transition occurs) and jump probabilities (the probability of transitioning to a state given that a transition occurs). If these quantities are of interest, a model parameterization which allows easily interpretable covariate effects to act on these quantities would be useful, especially as current parameterizations may not enable such interpretation. We consider this issue to motivate a modification of the original three-state model.

The next section introduces the psoriatic arthritis data on which this analysis is based.

2. Psoriatic arthritis data

Psoriatic arthritis is an inflammatory arthritis that is associated with the skin condition psoriasis. At the University of Toronto psoriatic arthritis clinic, over 1000 patients have been followed up longitudinally since it began in 1978 with clinic visits scheduled 6–12 months apart. In particular, at these visits, the active and damaged joint counts are recorded at the individual joint level, among other measurements, and therefore permit statistical modelling at this level of detail. In this investigation, focus will be on the 28 hand joints (14 joints in each hand; see Fig. 1 for more details), which can result in severe physical disability if active and/or damaged. Furthermore, this investigation is based on 743 patients who entered the clinic with no damage in either hand, so that patients are more comparable in their initial state of disease progression, and had more than two clinic visits. A dynamic covariate which requires previous observations will be constructed in the next section. Of this subset of patients, 69% (514 of 743 patients) had no damage at the end of their follow-up, which motivates consideration of a stayer population. The mean follow-up time was 10 years and 8 months with an interquartile range of 11 years and 6 months. The mean and median number of clinic visits were 12.7 and 8 respectively, and this ranged from 2 to 57. The mean and median intervisit times were 10 and 6 months, with a standard deviation of 1 year and 3 months. At clinic entry, the mean age at onset of arthritis was 36 years and 8 months with a standard deviation of 13 years and 4 months, whereas the mean duration of arthritis was 5 years and 2 months with a standard deviation of 7 years and 2 months. Furthermore, 55% of patients were male and 45% female.

In total, there were 264 208 observed transitions over all hand joints in the data. The observed transition matrix is
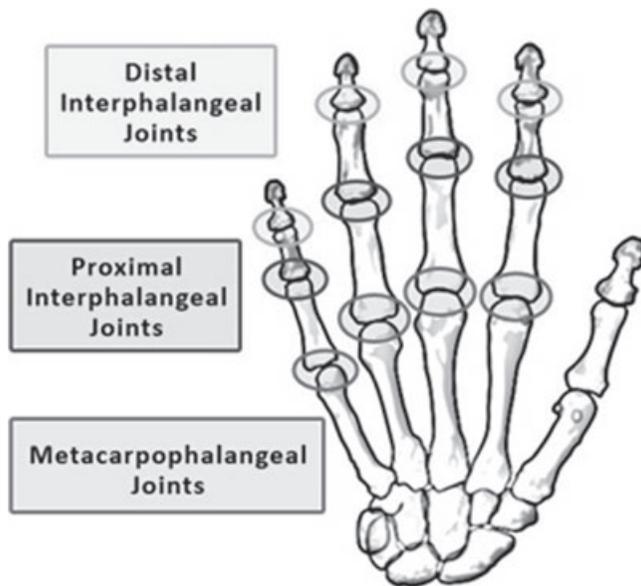
Fig. 1.   Diagram of the type of joints in each hand: this investigation focuses on 14 joints in each hand consisting of the distal interphalangeal, proximal interphalangeal and metacarpophalangeal joints in each finger and the proximal interphalangeal and metacarpophalangeal joints in the thumb (the figure was obtained from the arthritis fact sheet on the Georgia Tech Web page usability.gtri.gatech.edu)

$$
\begin{array}{cccc}
 & \bar{A}\bar{D} & A\bar{D} & \bar{A}D & AD \\
\begin{array}{c} \bar{A}\bar{D} \\ A\bar{D} \\ \bar{A}D \\ AD \end{array}
\left(\begin{array}{cccc}
217976 & 11008 & 623 & 158 \\
12250 & 8599 & 200 & 169 \\
0 & 0 & 10935 & 680 \\
0 & 0 & 882 & 728
\end{array}\right)
\end{array}
$$

where $\bar{A}$ and A denote the absence and presence of activity in the joint respectively, and $\bar{D}$ and D denote whether the joint has been clinically assessed as not damaged and damaged respectively.

The next section describes a six-state model which will be useful for jointly investigating the activity and damage processes.

3. Six-state model for transition intensities

Multistate models provide a convenient framework when the evolution of a stochastic process is of interest (Commenges, 1999; Anderson, 2002). This investigation demonstrates their use for the joint analysis of the disease activity and damage processes occurring in each individual hand joint. Specifically, consider the following four-state representation that is depicted in Fig. 2.

The process characteristics, particularly the reversibility of activity and permanent nature of damage, are reflected in the non-zero transition intensities which describe the instantaneous rate of transitioning between states. It is implicit that this representation describes the possible transitions of movers since λ_{13} and $\lambda_{24} > 0$. If, however, a stayer population exists with regard to developing damaged hand joints, their disease activity process can be described by the multistate diagram in Fig. 3.

Let λ_{rsij}^{l} denote the transition intensity from state r to s for the lth joint of the ith patient at the jth clinic visit (at each clinic visit all 28 hand joints are observed). Initially to investigate
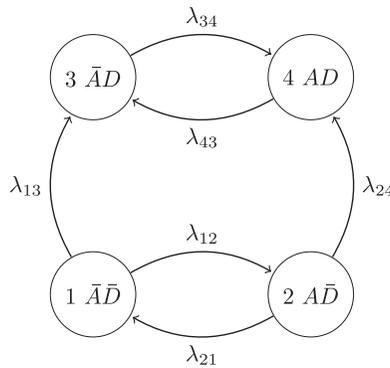

Fig. 2.   Multistate model describing the disease activity and damage processes jointly for movers
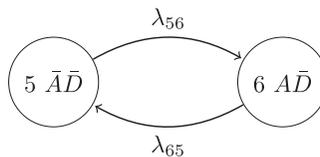

Fig. 3.   Multistate model describing the activity process for stayers

specific clinical aspects described in Section 1 (and to formulate a more parsimonious model), the transition intensities are parameterized as follows:

$$\left.\begin{aligned}
\lambda_{34ij}^{l} &= \lambda_{12ij}^{l} \exp(\beta_{\text{Damaged}}^{\bar{A}A}), \\
\lambda_{43ij}^{l} &= \lambda_{21ij}^{l} \exp(\beta_{\text{Damaged}}^{A\bar{A}}), \\
\lambda_{24ij}^{l} &= \lambda_{13ij}^{l} \exp(\beta_{\text{Active}}^{\bar{D}D}), \\
\lambda_{56ij}^{l} &= \lambda_{12ij}^{l} \exp(\beta_{\text{Stayer}}^{\bar{A}A}), \\
\lambda_{65ij}^{l} &= \lambda_{21ij}^{l} \exp(\beta_{\text{Stayer}}^{A\bar{A}}).
\end{aligned}\right\} \tag{1}$$

Thus regression coefficients are used to provide a simple representation of the main effect of damage and mover–stayer status on activity transitions as well as the main effect of activity on the damage transition. We note that more complex models involving interaction effects can be developed for estimating differential effects across various subgroups of patients. Furthermore, we let

$$\left.\begin{aligned}
\lambda_{12ij}^{l} &= \lambda_{0}^{\bar{A}A} \exp(\boldsymbol{\beta}^{\bar{A}A}\mathbf{z}_{ij}^{l} + u_{ij}), \\
\lambda_{21ij}^{l} &= \lambda_{0}^{A\bar{A}} \exp(\boldsymbol{\beta}^{A\bar{A}}\mathbf{z}_{ij}^{l} + \alpha u_{ij}), \\
\lambda_{13ij}^{l} &= \lambda_{0}^{\bar{D}D} \exp(\boldsymbol{\beta}^{\bar{D}D}\mathbf{z}_{ij}^{l} + v_{ij}),
\end{aligned}\right\} \tag{2}$$

where $\lambda_{0}^{\bar{A}A}$, $\lambda_{0}^{A\bar{A}}$ and $\lambda_{0}^{\bar{D}D}$ are constant baseline intensities, $\boldsymbol{\beta}^{\bar{A}A}$, $\boldsymbol{\beta}^{A\bar{A}}$ and $\boldsymbol{\beta}^{\bar{D}D}$ are vectors of regression coefficients, $\mathbf{z}_{ij}^{l}$ is a vector of covariates that are associated with the lth joint from the ith patient at the jth clinic visit, and u_{ij} and v_{ij} are realizations of zero-mean bivariate normal observation level random effects. Here $\alpha \in \mathbb{R}$ is an unknown parameter to be estimated which allows u_{ij} to act differently across the different transition intensities associated with the activity process. Although not formally stated, we include time-dependent dynamic covariates in $\mathbf{z}_{ij}^{l}$ to relax the Markov assumption. Specifically, the observed history of the activity process is summarized through a joint-specific covariate denoted as the adjusted mean activity AMA (Ibañez *et al.* (2003); Fig. 4 provides a description), whereas a patient's state of disease progression is reflected through the current number of damaged joints attained. On average, AMA was calculated as 0.093 with a standard deviation of 0.19 in our data. Given these dynamic covariates, current transition intensities from the multistate process are then assumed independent of previous process history, i.e. the Markov assumption. The random effects are assumed independent across time (with respect to j) and can be seen as accounting for unobserved heterogeneity not due to previous process history (where adjustments to unobserved heterogeneity related to previous history are provided through the dynamic covariates), which is still unaccounted for in the model. It is worth noting that the explicitly specified regression coefficients in expressions (1) and (2) correspond to covariates with different modelling assumptions. The covariates in $\mathbf{z}_{ij}^{l}$ are assumed to remain constant between clinic visits and therefore are relevant when this is true (time invariant covariates) or a reasonable approximation (e.g. when the covariate process is unlikely to be highly fluctuating between clinic visits). For simplicity, such covariates are also usually included if understanding the relationship between these covariates and the outcome is not of primary interest but some form of adjustment for these covariates is necessary. In contrast, the regression coefficients β_{Damaged} and β_{Active} are describing the effect of a binary variable representing a joint being damaged and active respectively while reflecting the stochastic nature of these processes and therefore provide more realistic measures of association. This is especially useful because these are the clinical aspects of primary interest. The regression coefficient β_{Stayer} is similar in nature to α as it describes the effect of a partially observable binary variable
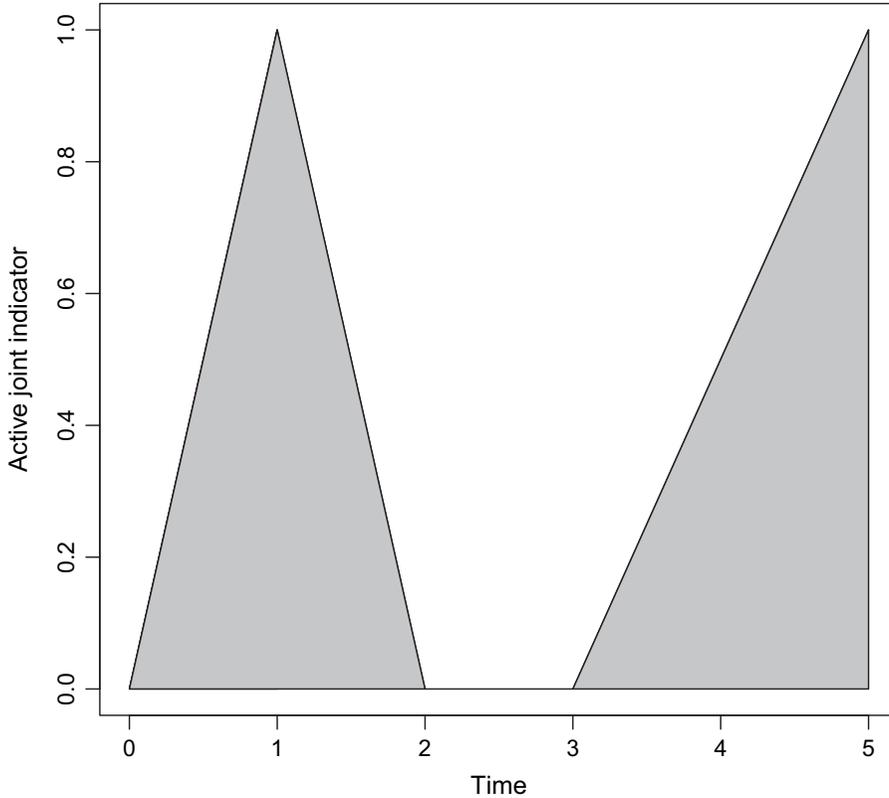
Fig. 4.   Let $x(t)$ be realized values of $X(t)$, a binary stochastic process describing the activity process of a joint: specifically at time t, $x(t) = 1$ corresponds to the joint being active and $x(t) = 0$ corresponds to the joint being not active; AMA(t) is then calculated as $(1/t) \int_0^t x(s) \, ds$, thus resulting in a bounded measure between [0, 1]; as $x(t)$ is intermittently observed and therefore the true path of $x(t)$ is not known, $\int_0^t x(s) \, ds$ is approximated as the area under the linearly interpolated observations of $x(t)$; for example, if a joint was observed at $t = (0, 1, 2, 3, 5)$ such that $x(t) = (0, 1, 0, 0, 1)$ respectively, then AMA(5) is approximated as the shaded area divided by 5, i.e. hence 0.4

(stayer $= 1$ and mover $= 0$); it can only be known that patients with damage are movers and that patients with no damage are either movers or stayers.

The motivation behind the use of observation level random effects arose from the potential need to allow for non-predictable changes in unobserved heterogeneity (Unkel *et al.*, 2014) with regard to the damage and activity processes. The current methodology in the literature primarily uses patient level random effects. This forces unobserved heterogeneity to be time invariant and we assess this assumption in our context by fitting this model (i.e. $\{U_{ij}, V_{ij}\} = \{U_i, V_i\} \; \forall \; j$) in addition to the model proposed. Likelihood values can then be used to compare informally the usefulness of the modelling framework proposed (models are non-nested but contain the same number of parameters; hence the same penalty terms are obtained if information criteria such as the Akaike information criterion are used). In general, the estimability of random-effects models, particularly variance components, will be driven by the random-effects structure and the variability of the data. The observation level random-effects structure incorporates fewer shared random effects between transition intensities than does its patient level counterpart. As a consequence, there will be less variability between transition intensities containing the same

random effect, thus making it more difficult to estimate the random-effect variance when using the observation level random-effects structure. This would especially be so when a single multistate process is of interest, where substantial heterogeneity in the data (which could be generated by constraining transition intensities) will probably be needed for observation level random-effects models to be estimable, whereas patient level random-effects models are likely to require considerably less heterogeneity (Satten (1999) and Cook (1999) considered a single multistate process with patient level random effects). In a clustered multistate process framework, heterogeneity is also generated through the differences across several multistate processes; thus the observation level random-effects structure is far more likely to be estimable, especially as the number of processes increases.

3.1. Maximum likelihood estimation

The model proposed is fitted by constructing and then maximizing the marginal likelihood. Let $X_i^l(t_{ij})$ denote the six-state process for the lth joint from the ith patient at time t_{ij}, where $\{t_{i1}, \ldots, t_{im_i}\}$ denotes the times of the jth clinic visit from the ith patient. Let C_i be a partially observable binary variable such that $C_i = 1$ with probability $1 - \pi_i$ if patient i is a mover (transitions are governed by the four-state model in Fig. 1), and $C_i = 0$ with probability π_i otherwise (transitions are governed by the two-state model in Fig. 2). Under the assumption that the conditional process (conditional on the piecewise constant dynamic covariates, $\{U_{ij} = u_{ij}, V_{ij} = v_{ij}\} \forall j$ and $C_i = c_i$) of the lth joint from the ith patient is Markov and time homogeneous (although the marginal process is not assumed to be Markov and/or time homogeneous), the conditional likelihood contribution from the lth joint of the ith patient is

$$\prod_{j=2}^{m_i-1} \mathbb{P}\{X_i^l(t_{ij+1}) = s_{ij+1}^l | X_i^l(t_{ij}) = s_{ij}^l; \mathbf{z}_{ij}^l, u_{ij}, v_{ij}, c_i\}$$

where s_{ij}^l represents the state corresponding to the specific combination of $(\{\bar{A}, A\}, \{\bar{D}, D\})$ observed at t_{ij} for the lth joint of the ith patient. More details on the likelihood construction for time homogeneous Markov models, particularly the form of the transition probabilities, can be found in Kalbfleisch and Lawless (1985). Appendix A provides the closed form transition probabilities of the six-state process. Here, for simplicity, the likelihood contribution from the process between t_{i1} and t_{i2} is excluded because AMA cannot be calculated at t_{i1}; it requires previous observations. As a diagnostic check of this simplification, we used activity at baseline to proxy AMA at baseline and this produced comparable results. If the assumption of independence between joints from the same patient is reasonable, conditional on the random effects U_{ij} and V_{ij}, then

$L_i(\mathbf{\Theta}|\mathbf{z}_{ij}, C_i)$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \prod_{l=1}^{28} \prod_{j=2}^{m_i-1} \mathbb{P}\{X_i^l(t_{ij+1}) = s_{ij+1}^l | X_i^l(t_{ij}) = s_{ij}^l; \mathbf{z}_{ij}^l, u_{ij}, v_{ij}, c_i\} \phi(u_{ij}, v_{ij}; \mathbf{0}, \Sigma) \mathrm{d}u_{ij} \mathrm{d}v_{ij}$$

represents the likelihood contribution from the ith patient, still conditional on the dynamic covariates and $C_i = c_i$. Here $\mathbf{\Theta}$ is a vector containing all the unknown parameters to be estimated (baseline intensities, regression coefficients and random-effects variance–covariance components) apart from the mover–stayer probabilities π_i, $\phi(u_{ij}, v_{ij}; \mathbf{0}, \Sigma)$ denotes the zero-mean bivariate normal density with covariance matrix Σ and $\mathbf{z}_{ij} = \{\mathbf{z}_{ij}^1, \ldots, \mathbf{z}_{ij}^{28}\}$. The overall marginal likelihood contribution from the ith patient is then

$$L_i(\mathbf{\Theta}^*|\mathbf{z}_{ij}) = \{(1 - \pi_i)L_i(\mathbf{\Theta}|\mathbf{z}_{ij}, C_i = 1)\}^{c_i^*} \{(1 - \pi_i)L_i(\mathbf{\Theta}|\mathbf{z}_{ij}, C_i = 1) + \pi_i L_i(\mathbf{\Theta}|\mathbf{z}_{ij}, C_i = 0)\}^{1-c_i^*}$$

with the overall marginal likelihood $L(\Theta^*|\mathbf{z}_{ij})$ obtained by taking the product of all likelihood contributions $L_i(\Theta^*|\mathbf{z}_{ij})$ from each patient. Here $\Theta^* = \{\Theta, \pi_i\}$ and c_i^* is a binary indicator such that $c_i^* = 1$ if damaged joints are observed from patient i at their last clinic visit and $c_i^* = 0$ otherwise. The bivariate numerical integrations were computed by firstly factorizing the bivariate density function into conditional densities, i.e. $\phi\{v_{ij}; \rho u_{ij}\sigma_v/\sigma_u, \sigma_v^2(1-\rho^2)\}\phi(u_{ij}; 0, \sigma_u^2)$, where σ_v^2 and σ_u^2 denote the respective variance components and ρ the correlation parameter, then using Gauss–Hermite quadrature to evaluate each integral with respect to u_{ij} and v_{ij} separately. The numbers of quadrature points for each integration were chosen to be 15 and 30 for the observation and patient level random-effects models respectively. Weights and nodes from the quadrature rule were then calculated using the R (R Core Team, 2015) package `statmod` (Smyth *et al.*, 2004). A sensitivity analysis indicated that further quadrature points provided negligible influence on parameter estimates and log-likelihood values. The log-likelihood was maximized using the BFGS (Broyden, 1970) optimization routine and asymptotic standard errors for parameter estimates were obtained by evaluating and then inverting the numerically derived Hessian matrix at the maximum likelihood estimates. The same estimation procedure was used in Yiu *et al.* (2017) to fit models that share the same essential characteristics with the models proposed in this paper. In their work, it was shown through simulation studies that, with a correctly specified model and sufficient follow-up information, the estimation procedure in this paper can provide reliable estimates of their model parameters.

The next subsection provides results of fitting the proposed model to the data described in Section 2.

3.2. Results

In addition to the aforementioned covariates, adjustment covariates for type of joint, presence of opposite or contralateral joint damage (the same joint in the opposite hand is damaged; opposite joint damaged equals 1 and 0 otherwise), sex (male $\equiv$ 1 and female $\equiv$ 0), age at onset of arthritis (in years) and duration of arthritis (in years) are provided through $\mathbf{z}_{ij}^l$. Joint type is represented through a five-level categorical variable with levels metacarpophalangeal, proximal interphalangeal, distal interphalangeal, thumb metacarpophalangeal and baseline thumb proximal interphalangeal. This covariate was included in the transition intensities that were associated with $\bar{A} \rightarrow A$ and $\bar{D} \rightarrow D$ with preliminary analysis demonstrating little evidence of differential recovery rates from activity (i.e. $A \rightarrow \bar{A}$). The binary variable specifying the presence of opposite joint damage is motivated by previous analyses (Cresswell and Farewell, 2011; O'Keeffe *et al.*, 2011) which indicate evidence of symmetric joint damage; the propensity of damage for a joint in a specific location to become damaged is increased if the contralateral joint in the other hand is earlier damaged.

Table 1 presents the results from fitting the model proposed (with dynamic covariates and an observation level random-effects structure) to the 743 psoriatic arthritis patients described in Section 2. For comparative purposes, a model with patient level random effects, i.e. $U_{ij} = U_i$ and $V_{ij} = V_i$, was also fitted. The results of this model and a comparison with the results of the model proposed are provided in Appendix B. The larger log-likelihood value of $-47\,389.11$ for the model proposed compared with the log-likelihood value of $-52\,784.84$ for the comparative model would suggest a preference for the model proposed.

3.2.1. Damage process

From Table 1, it is clear that both opposite joint damage and the number of damaged joints are strongly and positively associated with an increased damage progression rate. Thus this

Table 1. Parameter estimates and 95% Wald intervals resulting from fitting the six-state model (described in Section 3) to 743 psoriatic arthritis patients

Parameter	Estimates for the following transitions:		
	$\bar{A} \rightarrow A$	$A \rightarrow \bar{A}$	$\bar{D} \rightarrow D$
Damaged joint	-0.13 (-0.27, 0.0079)	-0.2 (-0.31, -0.088)	
Opposite joint damaged	0.17 (0.027, 0.3)	0.09 (-0.035, 0.21)	0.83 (0.6, 1.07)
Attained number of damaged joints	0.033 (0.021, 0.045)	-0.0039 (-0.014, 0.0061)	0.22 (0.18, 0.25)
Active joint			1.62 (1.3, 1.94)
AMA	2.72 (2.58, 2.87)	-0.49 (-0.61, -0.37)	2.01 (1.68, 2.34)
Metacarpophalangeal	0.3 (0.22, 0.372)		-0.84 (-1.10, -0.58)
Proximal interphalangeal	0.46 (0.38, 0.53)		-0.15 (-0.38, 0.089)
Distal interphalangeal	-0.18 (-0.26, -0.095)		0.49 (0.26, 0.73)
Thumb metacarpophalangeal	0.45 (0.36, 0.55)		0.45 (0.17, 0.72)
Sex	-0.69 (-0.79, -0.59)	0.017 (-0.055, 0.088)	0.2 (-0.047, 0.44)
Age at arthritis onset	0.0012 (-0.0031, 0.0055)	0.008 (0.0049, 0.011)	0.013 (0.0038, 0.023)
Arthritis duration	-0.021 (-0.027, -0.015)	0.0066 (0.0026, 0.011)	-0.01 (-0.023, 0.0028)
Stayer	1.99 (1.86, 2.12)	0.22 (0.11, 0.33)	
$\log(\lambda_0)$	-3.18 (-3.4, -2.95)	0.79 (0.63, 0.94)	-9.48 (-10.08, -8.89)
σ_u^2	2.07 (1.93, 2.21)		
α		-0.38 (-0.42, -0.35)	
σ_v^2			6.62 (5.89, 7.45)
ρ	0.16 (0.1, 0.21)		
π	0.14 (0.11, 0.18)		
Log-likelihood	-47389.11		

analysis supports the results in O'Keeffe *et al.* (2011) concerning symmetry even after adjusting for a greater number of process features, although not adjusting for the stochastic nature of the opposite joint damage process. Activity, both current (the joint is active while adjusting for its stochastic nature) and history (as described by AMA), is seen to be strongly and positively associated with damage progression. Regarding inference, the confidence interval for the regression coefficient that is associated with current activity is narrower than the corresponding interval that was reported in O'Keeffe *et al.* (2011). This has probably resulted from using updated data and a more efficient estimation procedure (through dynamic covariates and random effects as opposed to a working independence assumption with a robust covariance matrix adjustment). These results therefore provide greater confidence in the strong positive association between activity and damage, and implicitly strengthens the argument that was made regarding causality.

3.2.2. Activity process

There is evidence that the transition intensities that are associated with entering and leaving the active joint state reduces once a joint has become damaged. However, as the respective (95% Wald intervals) confidence intervals contain or are close to zero, this observation must currently be regarded as suggestive. The presence of opposite joint damage and the number of damaged joints seem to be moderately or weakly associated with the activity transition intensities. In particular little association is seen with transitioning from the active joint state to the not-active

joint state. The strong association between history of activity and current activity transition intensities is reassuring, since the interpretation of greater amounts of previous activity increasing the transition intensity to the active state while decreasing the transition intensity to the not-active state is intuitive.

3.2.3. Movers and stayers

The percentage of stayers ($100\pi\%$ where $\pi = \pi_i \,\forall i$) was estimated to be 14% (11%, 18%). Empirically, when compared with the 69% of patients who did not develop any damage, this estimate may seem to be a considerable underestimate of the true stayer proportion. However, because of the relationship between activity and damage, it is conceivable that many of these patients (those who did not develop damaged joints) did not develop damage because they were in the not-active state for long periods of continuous time, as opposed to being stayers *per se*. This observation is perhaps supported by Table 1, which suggests that movers have a vastly smaller transition intensity to the active state compared with stayers. A more specific investigation regarding the sojourn times of movers and stayers in the not-active joint state follows in the next section. It is also worth noting that the transition intensity to the not-active state is smaller for movers; however, it is far less pronounced.

As a diagnostic check for parameter estimation, we plot the profile log-likelihood for π in Fig. 5. From Fig. 5, it is clear that the numerical optimization routine converged at the maximum
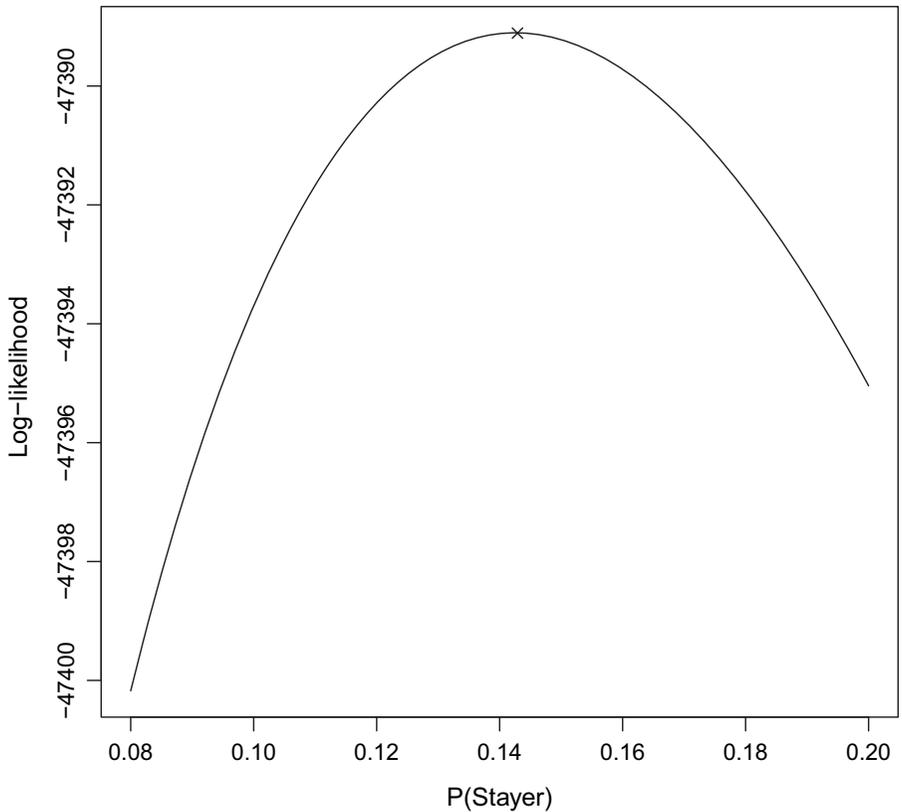
Fig. 5. Plot of the profile log-likelihood for π based on the six-state model in Section 3 (×, point at which the numerical optimization procedure converged)

of the profile log-likelihood for π. This indicates that the latent stayer proportion was identifiable under the assumptions of the six-state model in Section 3 because the profile log-likelihood has a quadratic shape.

4. Five-state model for mean sojourn times

In many settings, clinical interest lies in understanding the mean sojourn times from a particular time point (the mean amount of time that a process will spend in a state from a particular time point before a transition occurs) as a function of covariates at that particular time point. When two or more transitions are possible from a state of interest, the current methodology of investigating a covariate effect involves fixing other covariates at specified values (usually at their means or as a description of a particular patient) and then calculating the difference in mean sojourn times from a particular time point for that state by varying the covariate of interest. This methodology is implicit because, under the current multistate modelling parameterization in terms of transition intensities, a direct interpretation of covariate effects on the mean sojourn times is not straightforward when two or more transitions are possible from the state of interest; the mean sojourn time is a non-linear function of covariates from different transition intensities. This section considers the novel approach of modelling the mean sojourn times directly through a model reparameterization to obtain easily interpretable covariate effects on this quantity. In our situation, this approach is possible because there is a smooth bijection from the transition intensities to the mean sojourn times and jump probabilities (see the next paragraph). This implies that more elaborate but computationally intensive techniques such as the use of pseudo-observations (Anderson and Perme, 2010) can be avoided.

The specific context of interest concerns the sojourn times in the active and not-active states before damage. For simplicity, we can therefore revert to a three-state model for the movers by combining the activity process after damage has occurred into a single absorbing state, as depicted in Fig. 6. Similarly, the multistate diagram describing possible transitions for stayers is displayed in Fig. 7.

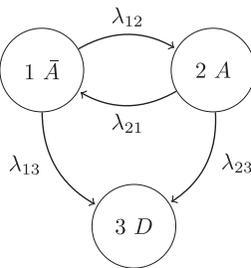Although continuous time Markov processes can be viewed in terms of transition intensities

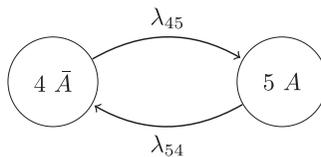**Fig. 6.** Multistate model describing activity and damage processes jointly for movers

Fig. 7. Multistate model describing the activity process for stayers

specifying risks of transitioning through states (as thought of previously), another natural way to view such processes is in terms of sojourn times and jump probabilities being associated with each state. As mentioned, the sojourn time in a state from a particular time point describes the amount of time that the process will spend in that state from that particular time point before a transition occurs. Then, at the point of transitioning, the jump probabilities inform the multinomial distribution of the possible set of states in which the process will jump to next. For the conditional process that we consider (a continuous time Markov process), the sojourn time in state r from the time of entry into state r can be shown to have an exponential distribution with mean $1/\Sigma_{k \neq r} \lambda_{rk}$ and jump probabilities given by $\mathbb{P}(\text{jumps to state } k \,|\, \text{jumps from state } r) = \lambda_{rk}/\Sigma_{k \neq r} \lambda_{rk}$. See pages 259–260 in Grimmett and Stirzaker (2001) for more details. Thus, by using the lack of memory property of the exponential distribution, we make the simplifying assumption that the sojourn times from a particular time point do not depend on the (unobserved) time that has already elapsed in that state conditional on time-dependent covariates and random effects at that particular time point and the time-independent mover–stayer status.

To investigate the difference in the activity process between movers and stayers again more simply, we parameterize as follows:

$$
\mu_{4ij}^{l} = \mu_{1ij}^{l} \exp(\beta_{\text{Stayer}}^{\bar{A}}),
$$
$$
\mu_{5ij}^{l} = \mu_{2ij}^{l} \exp(\beta_{\text{Stayer}}^{A}),
$$

where μ_{rij}^{l} represents the mean sojourn time in state r for the lth joint of the ith patient from time t_{ij}. Regression models for the mean sojourn times and jump probabilities can then be specified as follows:

$$
\left.
\begin{aligned}
\mu_{1ij}^{l} &= \mu_0^{\bar{A}} \exp(\boldsymbol{\beta}^{\bar{A}} \mathbf{z}_{ij}^{l} + u_{ij}), \\
\mu_{2ij}^{l} &= \mu_0^{A} \exp(\boldsymbol{\beta}^{A} \mathbf{z}_{ij}^{l} + \alpha_1 u_{ij}), \\
p_{13ij}^{l}/(1 - p_{13ij}^{l}) &= p_0^{\bar{A}D} \exp(\boldsymbol{\beta}^{\bar{A}D} \mathbf{z}_{ij}^{l} + v_{ij}), \\
p_{23ij}^{l}/(1 - p_{23ij}^{l}) &= p_0^{AD} \exp(\boldsymbol{\beta}^{AD} \mathbf{z}_{ij}^{l} + \alpha_2 v_{ij}),
\end{aligned}
\right\}
\tag{3}
$$
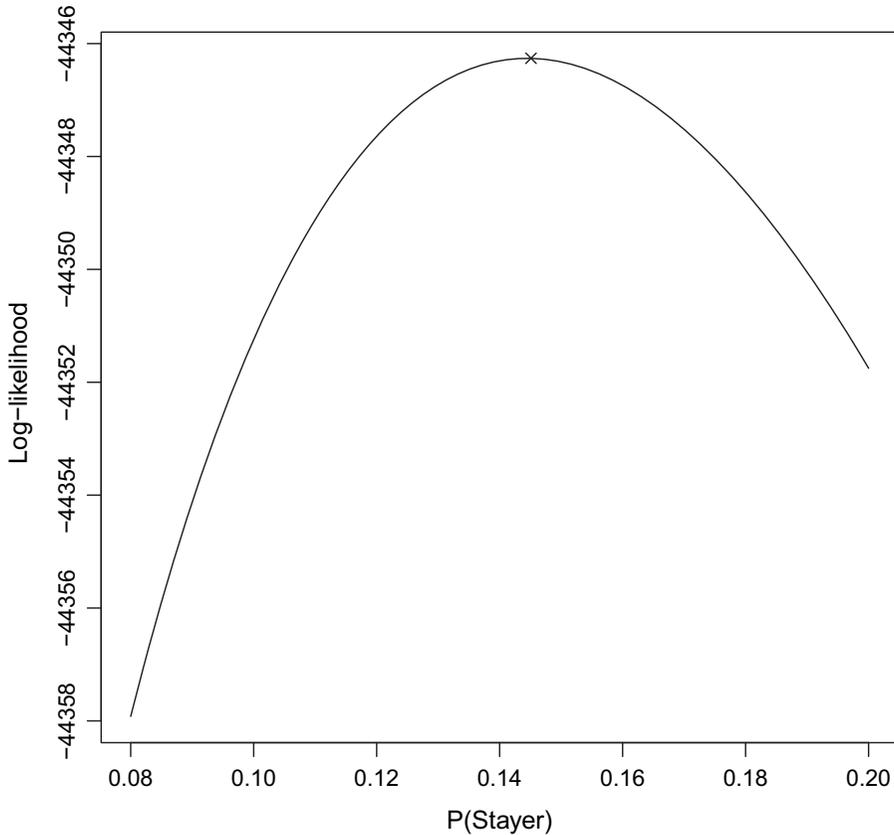
where $p_{\bar{A}Dij}^{l}$ and p_{ADij}^{l} denote the jump probabilities from state $\bar{A} \to D$ and $A \to D$ respectively for the lth joint of the ith patient at t_{ij}. The right-hand side of each regression equation in expression (3) contains a baseline (indicated by 0 in the subscript) multiplied by the exponent of the sum of a linear predictor and linear function of realizations of random effects, as before. Dynamic covariates (AMA and the attained number of damaged joints) and random effects are again included to reflect features of the processes that were described in Section 3. The random effects follow a zero-mean bivariate normal distribution and are independent across time.

The model fitting procedure follows from Section 3.1 after having specified the transition probabilities which as a function of transition intensities can be found in Appendix C. Thus the following set of equations completes the procedure:

$$
\lambda_{12} = (1 - p_{13})/\mu_1,
$$
$$
\lambda_{13} = p_{13}/\mu_1,
$$
$$
\lambda_{21} = (1 - p_{23})/\mu_2,
$$
$$
\lambda_{23} = p_{23}/\mu_2,
$$
$$
\lambda_{45} = 1/\mu_4,
$$
$$
\lambda_{54} = 1/\mu_5.
$$

Table 2. Parameter estimates and 95% Wald intervals resulting from fitting the five-state model (described in Section 4) to 743 psoriatic arthritis patients

Parameter	Sojourn times		Jump probabilities	
	$\bar{A}$	A	$\bar{A} \to D$	$A \to D$
Opposite joint damaged	−0.28 (−0.44, −0.12)	−0.1 (−0.27, 0.061)	1.6 (0.87, 2.34)	0.54 (−0.18, 1.26)
Attained number of damaged joints	−0.09 (−0.11, −0.074)	0.0075 (−0.0057, 0.021)	0.21 (0.14, 0.29)	0.16 (0.11, 0.22)
AMA	−2.98 (−3.13, −2.83)	0.39 (0.26, 0.52)	−0.25 (−1.02, 0.52)	0.84 (0.21, 1.48)
Sex	0.68 (0.58, 0.78)	0.0053 (−0.067, 0.078)	1.24 (0.74, 1.74)	0.43 (−0.022, 0.89)
Age at arthritis onset	−0.0021 (−0.0064, 0.0023)	−0.0078 (−0.011, −0.0047)	1.24 (0.74, 1.74)	0.43 (−0.022, 0.89)
Arthritis duration	0.019 (0.014, 0.025)	−0.0036 (−0.0076, 0.00047)	0.037 (0.016, 0.059)	−0.023 (−0.051, 0.0053)
Stayer	−1.93 (−2.06, −1.8)	−0.17 (−0.29, −0.049)		
$\log(\mu_0)$	3.04 (2.82, 3.25)	−0.83 (−0.99, −0.67)		
$\log(p_0)$			−9.27 (−10.9, −7.65)	−5.68 (−6.81, −4.55)
σ_u^2	2.05 (1.92, 2.2)			
α_1		−0.35 (−0.39, −0.32)		
σ_v^2			17.4 (12.66, 23.92)	
α_2				0.44 (0.32, 0.56)
ρ	0.0041 (−0.048, 0.056)			
π	0.15 (0.12, 0.18)			
Log-likelihood	−44346.26			

Fig. 8. Plot of the profile log-likelihood for π based on the five-state model in Section 4 (×, point at which the numerical optimization procedure converged)

4.1. Results

Along with dynamic covariates, the presence of opposite joint damage, sex, age at onset of arthritis and duration of arthritis were included in the analysis, as before. Table 2 presents the results from fitting the five-state model that was described in Section 4 to the 743 psoriatic arthritis patients described in Section 2. From Table 2, the presence of opposite joint damage provides a slight increase in the mean sojourn time in the not-active joint state and greatly increases the probability of directly transitioning to damage (as opposed to active and not damaged) once a transition occurs. However, when a joint is active, there is little evidence to suggest that opposite joint damage influences the sojourn times nor the next state probability. These results indicate that the presence of opposite joint damage is particularly relevant when a joint is not active. A large number of damaged joints, although substantially increasing the jump probabilities to damage as opposed to active or not active, provides little effect on the mean sojourn times in the active and not-active states. As expected, greater amounts of previous activity (as described by AMA) decrease the mean sojourn time in the not-active state and increase the mean sojourn time in the active state. Table 2 also suggests that AMA is strongly and positively associated with the jump probability to damage when in the active state but not in the not-active state. Thus current activity is strengthened by the history of activity when dictating the next state of the process, but jumping to damage or

Table 3. Parameter estimates and 95% Wald intervals resulting from fitting the five-state model with a patient level random-effects structure to 743 psoriatic arthritis patients

Parameter	Sojourn times		Jump probabilities	
	$\bar{A}$	A	$\bar{A} \rightarrow D$	$A \rightarrow D$
Opposite joint damaged	−0.28 (−0.44, −0.11)	−0.19 (−0.35, −0.026)	0.95 (0.44, 1.46)	0.44 (0.01, 0.86)
Attained number of damaged joints	0.012 (−0.0022, 0.026)	−0.033 (−0.044, −0.022)	−0.3 (−0.45, −0.14)	0.056 (0.019, 0.093)
AMA	−1.98 (−2.11, −1.85)	0.29 (0.17, 0.4)	0.7 (0.068, 1.32)	0.87 (0.34, 1.4)
Sex	0.79 (0.7, 0.88)	0.012 (−0.054, 0.078)	1.14 (0.51, 1.77)	0.36 (−0.066, 0.79)
Age at arthritis onset	−0.014 (−0.019, −0.0098)	−0.0028 (−0.0054, −0.00025)	0.032 (0.0095, 0.054)	−0.0082 (−0.03, 0.013)
Arthritis duration	0.036 (0.032, 0.039)	−0.0024 (−0.0062, 0.0013)	0.18 (0.13, 0.23)	0.041 (0.017, 0.065)
Stayer	−2.23 (−2.36, −2.1)	−0.35 (−0.46, −0.23)		
$\log(\mu_0)$	2.71 (2.49, 2.94)	−0.91 (−1.04, −0.77)		
$\log(P_0)$			−8.8 (−10.43, −7.18)	−5.61 (−6.59, −4.62)
σ_u^2	1.28 (1.16, 1.41)			
α_1		−0.24 (−0.29, −0.19)		
σ_v^2			6.96 (4.41, 10.98)	
α_2				0.63 (0.45, 0.81)
ρ	−0.057 (−0.15, 0.032)			
π	0.16 (0.13, 0.2)			
Log-likelihood	−49256.31			

active from the not-active state could be unaffected by the history of activity. As hypothesized in the previous section, stayers have a far shorter mean sojourn time in the not-active state, and a slightly shorter sojourn time in the active state. Finally, the estimated stayer proportion in Table 2 is approximately equivalent to the stayer proportion that was reported in Table 1. Under the assumptions of the five-state model, the numerical optimization procedure could identify the parameter π as it converged at the maximum of its profile log-likelihood (Fig. 8).

Table 3 provides the results of fitting the five-state model with a patient level random-effects structure (i.e. $U_{ij} = U_i$ and $V_{ij} = V_i \; \forall \; j$) to the data that were described in Section 2. The resulting log-likelihood value was $-49\,256.31$, which is far smaller than $-44\,346.26$ obtained from the model proposed. In this context, the observation level random-effects structure, after including dynamic covariates, again seems to be the more appropriate random-effects structure for the data.

5. Discussion

This research was motivated by reproducing prior results with an updated data set and undertaking new investigations into disease course and progression. For this, a single unifying clustered multistate modelling framework which allows simultaneous investigations of multiple clinical aspects was proposed. The results obtained therefore yield greater confidence when compared with multiple univariate investigations, which were performed previously, since they are based on adjusting for other important process characteristics. From a clinical perspective, the relationship between activity and damage was demonstrated as pronounced since both history and current activity were positively related to damage progression and jumping to damage once a joint immediately leaves the active state. In terms of the reverse relationship, the onset of damage is seen to slow the activity process although the confidence intervals for the relevant regression coefficients indicate that no change is a distinct possibility, maybe because of far fewer observed transitions after damaged has occurred. Interestingly, both models seem to identify a subpopulation of approximately 15% who are rapidly fluctuating in their activity process yet are at minimal risk of damage, perhaps because they have shorter sojourn times in the active joint state. An avenue of future work could involve identifying these patients especially because their treatment strategies should conceivably not consist of potent drugs, which may cause unpleasant side effects, but soft drugs to reduce joint swelling and pain. It is also reassuring that neither model contradicts any well-held clinical beliefs.

From a statistical point of view, the novel aspects of this research include the proposing of an observation level random-effects structure combined with dynamic covariates, a mover–stayer structure whereby movers and stayers can have different effects on transition intensities in which they are not implicitly defined for and a model parameterization which allows easily interpretable covariate effects to act on the sojourn times and jump probabilities. In our context, the usefulness of the methodology proposed was demonstrated through new clinical insights and substantial improvements in likelihood values over the use of standard methodology (patient level random-effect models). Although the methodology proposed was described in terms of specific, but fairly complex, six- and five-state models for the margins, extensions to general clustered continuous time Markov models are straightforward. In particular, the proposed model parameterization in terms of sojourn times and jump probabilities, and mover–stayer effects on transition intensities are also applicable to univariate Markov multistate processes and therefore can provide a useful framework for inference in many clinical settings.

Overall, this research represents our efforts to provide a comprehensive investigation into

many clinical aspects of interest at the finest level of detail (individual joint level). Although there are foreseeable model extensions, it is important to bear in mind the computationally intensive nature of fitting clustered multistate models with random effects, especially when reversible multistate models are involved (transitions to and from states exist). Some examples of potentially more appropriate extensions could include relaxing the time homogeneity assumption beyond adjusting for duration of arthritis, relating previous history to current transitions through more accurate measures than the proposed dynamic covariates, e.g. by specifying a flexible correlation structure for the observation level random effects, and dividing the damage onset state into various states of severity of damage. Such extensions, as with many others, will usually require transitioning from piecewise constant approximations to reflecting the true stochastic nature of the outcome, covariates and latent processes, which has been seen as one of the main drivers of making the model fitting procedure more complex, because of the larger number of integrations or differential equations that are required to be computed or solved. Nevertheless, as demonstrated here, it is important to identify and provide adjustments for important process characteristics where possible, in which model comparison is useful. With respect to understanding clinical aspects of psoriatic arthritis, this has provided new knowledge and greater confidence in prior results based on less general methodology.

Acknowledgements

Appendix A

The transition probabilities, i.e. $p_{rs}(t) \equiv p_{rs}(t; \boldsymbol{\lambda}) = \mathbb{P}\{X(t+t_0) = s | X(t_0) = r; \boldsymbol{\lambda}\} \ \forall t_0 \geqslant 0$ where $\boldsymbol{\lambda}$ is a vector of the required transition intensities, for the four- and two-state process depicted in Figs 2 and 3 respectively can be calculated by evaluating the matrix exponential of the relevant transition intensity matrix. Specifically, let

$$Q_1 = \begin{pmatrix} -\lambda_{12} - \lambda_{13} & \lambda_{12} & \lambda_{13} & 0 \\ \lambda_{21} & -\lambda_{21} - \lambda_{24} & 0 & \lambda_{24} \\ 0 & 0 & -\lambda_{34} & \lambda_{34} \\ 0 & 0 & \lambda_{43} & -\lambda_{43} \end{pmatrix};$$

then the transition probabilities for the four-state process can be obtained from the (r, s)th entry of $\exp(Q_1 t)$. We use Mathematica to compute the following matrix exponential and this results in

$$p_{11}(t) = \frac{\exp(-\Lambda t/2)}{2\gamma}\{(\Lambda_1 + \gamma)\exp(-\gamma t/2) + (\gamma - \Lambda_1)\exp(\gamma t/2)\},$$

$$p_{12}(t) = \frac{\lambda_{12}}{\gamma_1}[\exp\{-(\Lambda - \gamma_1)t/2\} - \exp\{-(\Lambda + \gamma_1)t/2\}],$$

$$
\begin{aligned}
p_{13}(t) = {} & \frac{\lambda_{43}}{\Lambda_2} + \frac{\exp(-\Lambda_2 t)}{\Lambda_2}\frac{\lambda_{13}\lambda_{34}(\lambda_{21} + \lambda_{24} - \Lambda_2) - \lambda_{12}\lambda_{24}\lambda_{43}}{(\lambda_{21} + \lambda_{24} - \Lambda_2)(\lambda_{13} - \Lambda_2) - \lambda_{12}(\lambda_{24} - \Lambda_2)} \\
& + \lambda_{13}\exp\{-(\Lambda - \gamma)t/2\}\frac{\lambda_{12}\lambda_{21} + (\lambda_{12} + \lambda_{13})(\Lambda_1 - \gamma)/2 + \lambda_{43}(-\Lambda_1 + \gamma - 2\lambda_{34})/2 + \gamma_2}{-(\Lambda - \gamma)^3/2 + 3(\Lambda - \gamma)^2(\Lambda + \Lambda_2)/4 - (\Lambda - \gamma)\gamma_3 + \gamma_4} \\
& + \lambda_{13}\exp\{-(\Lambda + \gamma)t/2\}\frac{\lambda_{12}\lambda_{21} + (\lambda_{12} + \lambda_{13})(\Lambda_1 + \gamma)/2 + \lambda_{43}(-\Lambda_1 - \gamma - 2\lambda_{34})/2 + \gamma_2}{-(\Lambda + \gamma)^3/2 + 3(\Lambda + \gamma)^2(\Lambda + \Lambda_2)/4 - (\Lambda + \gamma)\gamma_3 + \gamma_4},
\end{aligned}
$$

$$p_{14}(t) = 1 - p_{11}(t) - p_{12}(t) - p_{13}(t),$$

where

$$\Lambda = \lambda_{12} + \lambda_{13} + \lambda_{21} + \lambda_{24},$$
$$\Lambda_1 = \lambda_{12} + \lambda_{13} - \lambda_{21} - \lambda_{24},$$
$$\Lambda_2 = \lambda_{34} + \lambda_{43},$$
$$\gamma = \sqrt{[\Lambda^2 - 4\{\lambda_{12}\lambda_{24} + \lambda_{13}(\lambda_{21} + \lambda_{24})\}]},$$
$$\gamma_1 = \sqrt{\{\lambda_{12}^2 + 2\lambda_{12}(\lambda_{13} + \lambda_{21} - \lambda_{24}) + (-\lambda_{13} + \lambda_{21} + \lambda_{24})^2\}},$$
$$\gamma_2 = \frac{\lambda_{43}}{\lambda_{13}}(\lambda_{12}\lambda_{24} + \lambda_{13}\lambda_{34}),$$
$$\gamma_3 = \Lambda_2(\lambda_{21} + \lambda_{24}) + \lambda_{13}\lambda_{21} + (\lambda_{12} + \lambda_{13})(\lambda_{24} + \Lambda_2),$$
$$\gamma_4 = \Lambda_2\{\lambda_{13}\lambda_{21} + \lambda_{24}(\lambda_{12} + \lambda_{13})\}.$$

As the four-state process is symmetric, it is easy to verify that

$$p_{21}(t; \lambda_{12}, \lambda_{13}, \lambda_{21}, \lambda_{24}, \lambda_{34}, \lambda_{43}) = p_{12}(t; \lambda_{21}, \lambda_{24}, \lambda_{12}, \lambda_{13}, \lambda_{43}, \lambda_{34}),$$
$$p_{22}(t; \lambda_{12}, \lambda_{13}, \lambda_{21}, \lambda_{24}, \lambda_{34}, \lambda_{43}) = p_{11}(t; \lambda_{21}, \lambda_{24}, \lambda_{12}, \lambda_{13}, \lambda_{43}, \lambda_{34}),$$
$$p_{23}(t; \lambda_{12}, \lambda_{13}, \lambda_{21}, \lambda_{24}, \lambda_{34}, \lambda_{43}) = p_{14}(t; \lambda_{21}, \lambda_{24}, \lambda_{12}, \lambda_{13}, \lambda_{43}, \lambda_{34}),$$
$$p_{24}(t; \lambda_{12}, \lambda_{13}, \lambda_{21}, \lambda_{24}, \lambda_{34}, \lambda_{43}) = p_{13}(t; \lambda_{21}, \lambda_{24}, \lambda_{12}, \lambda_{13}, \lambda_{43}, \lambda_{34}).$$

Finally, we have

$$p_{33}(t) = 1 - p_{34}(t),$$
$$p_{34}(t) = \lambda_{34}\{1 - \exp(-\Lambda_2 t)\}/\Lambda_2,$$
$$p_{43}(t) = \lambda_{43}\{1 - \exp(-\Lambda_2 t)\}/\Lambda_2,$$
$$p_{44}(t) = 1 - p_{34}(t),$$
$$p_{55}(t) = 1 - p_{56}(t),$$
$$p_{56}(t) = \lambda_{56}\{1 - \exp(-\Lambda_3 t)\}/\Lambda_3,$$
$$p_{65}(t) = \lambda_{65}\{1 - \exp(-\Lambda_3 t)\}/\Lambda_3,$$
$$p_{66}(t) = 1 - p_{56}(t),$$

where $\Lambda_3 = \lambda_{56} + \lambda_{65}$.

Appendix B

Table 4 provides the results from fitting the six-state model with a patient level random-effects structure to the 743 psoriatic arthritis patients that was described in Section 2. Rather reassuringly, most regression coefficients of primary interest, including the stayer proportion estimate, from Table 4 are seen to have similar interpretations to those obtained from Table 1. The regression coefficients that are associated with the dynamic covariates AMA and attained number of damaged joints have, however, resulted in markedly different estimates from those seen in Table 1. Specifically, the effects of the dynamic covariates are greatly attenuated and, furthermore, the attained number of damaged joints is now seen to have a possibly counterintuitive negative association with damage progression. Both dynamic covariates and patient level random effects adjust for a patient's propensity to gain or recover from activity and to gain damage, and thus are likely to be confounded when introduced simultaneously in the model. The regression coefficients that are associated with the dynamic covariates are now perhaps more difficult to interpret than before. Aalen *et al.* (2008) provides a discussion on the relationship between dynamic models (with dynamic covariates) and frailty models (with patient level random effects).

Appendix C

The transition probabilities of the three-state model were again calculated using the matrix exponential function in Mathematica. This resulted in

Table 4. Parameter estimates and 95% Wald intervals resulting from fitting the six-state model with a patient level random-effects structure to 743 psoriatic arthritis patients

	Results for the following transitions:		
	$\bar{A} \to A$	$A \to \bar{A}$	$\bar{D} \to D$
Damaged joint	−0.046 (−0.18, 0.086)	−0.18 (−0.27, −0.069)	
Opposite joint damaged	0.13 (−0.0031, 0.27)	0.087 (−0.037, 0.21)	0.72 (0.52, 0.92)
Attained number of damaged joints	−0.04 (−0.049, −0.03)	0.022 (0.013, 0.031)	−0.036 (−0.055, −0.018)
Active joint			1.32 (1.03, 1.61)
AMA	1.8 (1.68, 1.93)	−0.34 (−0.45, −0.23)	1.7 (1.39, 2)
Metacarpophalangeal	0.23 (0.16, 0.3)		−0.81 (−1.1, −0.57)
Proximal interphalangeal	0.36 (0.29, 0.43)		−0.18 (−0.4, 0.049)
Distal interphalangeal	−0.19 (−0.27, −0.12)		0.36 (0.14, 0.58)
Thumb metacarpophalangeal	0.39 (0.3, 0.48)		0.36 (0.095, 0.62)
Sex	−0.73 (−0.8, −0.66)	0.011 (−0.051, 0.073)	−0.2 (−0.45, 0.042)
Age at arthritis onset	0.014 (0.011, 0.018)	0.0042 (0.0015, 0.0069)	−0.012 (−0.024, −0.00034)
Arthritis duration	−0.037 (−0.04, −0.033)	0.0058 (0.0022, 0.0095)	0.057 (0.047, 0.067)
Stayer	2.25 (2.12, 2.37)	0.37 (0.26, 0.47)	
$\log(\lambda_0)$	−2.88 (−3.07, −2.69)	0.8 (0.66, 0.94)	−7 (−7.55, −6.45)
σ_u^2	1.33 (1.21, 1.46)		
α		−0.23 (−0.28, −0.19)	
σ_v^2			3.19 (2.73, 3.73)
ρ	0.28 (0.2, 0.35)		
π	0.17 (0.13, 0.21)		
Log-likelihood	−52784.84		

$$p_{11}(t) = \frac{x_1 \exp(r_1 t) - x_2 \exp(r_2 t)}{x_1 - x_2},$$

$$p_{12}(t) = \frac{x_1 x_2 \{\exp(r_2 t) - \exp(r_1 t)\}}{x_1 - x_2},$$

$$p_{13}(t) = 1 - p_{11}(t) - p_{12}(t),$$

$$p_{21}(t) = \frac{\exp(r_1 t) - \exp(r_2 t)}{x_1 - x_2},$$

$$p_{22}(t) = \frac{x_1 \exp(r_2 t) - x_2 \exp(r_1 t)}{x_1 - x_2},$$

$$p_{23}(t) = 1 - p_{21}(t) - p_{22}(t),$$

$$p_{3j}(t) = \mathbf{1}_{j=3},$$

where $\mathbf{1}_{j=3}$ is an indicator function taking the value 1 when $j = 3$ and 0 otherwise, and

$$r_1 = \frac{\lambda_{11} + \lambda_{22} + \sqrt{\{(\lambda_{11} - \lambda_{22})^2 + 4\lambda_{12}\lambda_{21}\}}}{2},$$

$$r_2 = \frac{\lambda_{11} + \lambda_{22} - \sqrt{\{(\lambda_{11} - \lambda_{22})^2 + 4\lambda_{12}\lambda_{21}\}}}{2},$$

$$x_j = \frac{r_j - \lambda_{22}}{\lambda_{21}}.$$

Here $\lambda_{11} = -\lambda_{12} - \lambda_{13}$ and $\lambda_{22} = -\lambda_{21} - \lambda_{23}$.

References

Aalen, O. O., Borgan, Ø. and Gjessing, H. K. (2008) *Survival and Event History Analysis: a Process Point of View*. New York: Springer.

Anderson, P. K. (2002) Multi-state models for event history analysis. *Statist. Meth. Med. Res.*, **11**, 91–115.

Anderson, P. K. and Perme, M. P. (2010) Pseudo-observations in survival analysis. *Statist. Meth. Med. Res.*, **19**, 71–99.

Broyden, C. G. (1970) The convergence of a class of double-rank minimisation algorithms. *J. Inst. Math. Appl.*, **6**, 76–90.

Commenges, D. (1999) Multi-state models in epidemiology. *Liftim. Data Anal.*, **5**, 315–327.

Cook, R. J. (1999) A mixed model for Markov processes under panel observation. *Biometrics*, **55**, 178–183.

Cook, R. J., Yi, G. Y., Lee, K. and Gladman, D. D. (2004) A conditional Markov model for clustered progressive multistate processes under incomplete observation. *Biometrics*, **60**, 436–443.

Cresswell, L. and Farewell, V. T. (2011) Assessment of joint symmetry in arthritis. *Statist. Med.*, **30**, 973–983.

Diao, L. and Cook, R. J. (2014) Composite likelihood for joint analysis of multiple multistate processes via copulas. *Biostatistics*, **15**, 690–705.

Frydman, H. (1984) Maximum likelihood estimation in the mover-stayer model. *J. Am. Statist. Ass.*, **79**, 632–638.

Grimmett, G. and Stirzaker, D. (2001) *Probability and Random Processes*. Oxford: Oxford University Press.

Ibañez, D., Urowitz, M. B. and Gladman, D. D. (2003) Summarizing disease features over time: I, Adjusted mean SLEDAI derivation and application to an index of disease activity in lupus. *J. Rheum.*, **9**, 1977–1982.

Kalbfleisch, J. D. and Lawless, J. F. (1985) The analysis of panel data under a Markov assumption. *J. Am. Statist. Ass.*, **80**, 863–871.

Lee, E. W. and Kim, M. Y. (1998) The analysis of correlated panel data using a continuous time Markov model. *Biometrics*, **54**, 1638–1644.

O'Keeffe, A. G., Tom, B. D. M. and Farewell, V. T. (2011) A case-study in the clinical epidemiology of psoriatic arthritis: multistate models and causal arguments. *Appl. Statist.*, **60**, 675–699.

O'Keeffe, A. G., Tom, B. D. M. and Farewell, V. T. (2012) Mixture distributions in multi-state modelling: some considerations in a study of psoriatic arthritis. *Statist. Med.*, **32**, 600–619.

R Core Team (2015) *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Satten, G. A. (1999) Estimating the extent of tracking in interval-censored chain-of-events data. *Biometrics*, **55**, 1228–1231.

Smyth, G., Hu, Y., Dunn, P., Phipson, B. and Chen, Y. (2004) {statmod}:statistical modeling. *R Package Version 1.4.20*. Walter and Eliza Hall Institute of Medical Research, Parkville. (Available from http://CRAN.R-project.org/package=statmod.)

Sutradhar, R. and Cook, R. J. (2008) Analysis of interval-censored data from clustered multistate processes: application to joint damage in psoriatic arthritis. *Appl. Statist.*, **57**, 553–566.

Tom, B. D. M. and Farewell, V. T. (2011) Intermittent observation of time-dependent explanatory variables: a multi-state modelling approach. *Statist. Med.*, **30**, 3520–3531.

Unkel, S., Farrington, C. P., Whitaker, H. J. and Pebody, R. (2014) Time varying frailty models and the estimation of heterogeneities in transmission of infectious diseases. *Appl. Statist.*, **63**, 141–158.

Yiu, S., Farewell, V. T. and Tom, B. D. M. (2017) Exploring the existence of a stayer population with mover–stayer counting process models: application to joint damage in psoriatic arthritis. *Appl. Statist.*, **66**, 669–690.