Web-based Supplementary Materials for "Joint latent class model for longitudinal data and interval-censored semi-competing events: Application to dementia"

Anaïs Rouanet\textsuperscript{1,2}, Pierre Joly\textsuperscript{1,2}, Jean-François Dartigues\textsuperscript{1,2}, Cécile Proust-Lima\textsuperscript{1,2} and Hélène Jacqmin-Gadda\textsuperscript{1,2}

\textsuperscript{1}INSERM, Centre INSERM U1912 - Epidemiologie - Biostatistiques, F-33076 Bordeaux, France
\textsuperscript{2}Université de Bordeaux, ISPED, 146 rue Léo Saignat, F-33076 Bordeaux, France

Contents

A Conditional log-likelihood of a semi-markovian model 1

B Application of the joint latent class illness-death model on PAQUID, adjusting on ApoE4 1

A Conditional log-likelihood of a semi-markovian model

In the case of a semi-markovian model, the individual conditional contribution to the log-likelihood of the multi-state model is changed for the observation patterns 1, 2, 5 and 6 as follows:

• Subject diagnosed with dementia (cases 1 and 2, Figure 2):

\[
P_{ig}^{(d)}(T_0, L_i, R_i, 1, T_i, \delta_i^D; \theta_G) = \int_{L_i}^{R_i} e^{-A_{01g}(u) - A_{02g}(u)} \alpha_{01g}(u) e^{-A_{12g}(T_i - u)} \alpha_{12g}(T_i - u)^{\delta_i^D} du
\]

• Subject with unknown dementia status at the end of follow-up \(T_i\) (cases 5 and 6, Figure 2):

\[
P_{ig}^{(u)}(T_0, L_i, R_i, 0, T_i, \delta_i^D; \theta_G) = e^{-A_{01g}(T_i) - A_{02g}(T_i)} \alpha_{02g}(T_i)^{\delta_i^D} + \int_{L_i}^{T_i} e^{-A_{01g}(u) - A_{02g}(u)} \alpha_{01g}(u) e^{-A_{12g}(T_i - u)} \alpha_{12g}(T_i - u)^{\delta_i^D} du
\]

B Application of the joint latent class illness-death model on PAQUID, adjusting on ApoE4

The model was applied on the sub-sample of subjects with ApoE4 measurement available. The sample under study involved only 619 subjects, who accepted blood testing for ApoE4 measurement, including 42.7% of men and 22.8% of ApoE4 carriers. These subjects were younger (mean age at entry: 73.9 vs 75.3, p-value < 0.001) and more educated (72.5% vs 64.4% with a high educational level, p-value < 0.001) than the sample with no
ApoE4 measurement. A total of 31.3% of subjects were diagnosed with dementia, including 27.3% who died during the follow-up, and 58.8% died before the dementia diagnosis.

The three transition intensities of the illness-death sub-model were adjusted on ApoE4, as well as the change-point mixed sub-model (intercept and slope), with all these regression parameters being common to the classes. The BIC of the models from 1 to 5 classes were respectively 25252, 25161, 25164, 25168 and 25216. Web Table 9 presents the estimated parameters for the model with 2 latent classes that had the lowest BIC. Education is associated with a lower transition intensity to dementia but not to death among healthy subjects. Gender is associated with lower transition intensities to death and ApoE4 carriers have a higher intensity transition to dementia.

These results are in line with previous analyses on PAQUID data which found that ApoE4 was not associated with cognitive level after adjustment on educational level and found no association with cognitive decline (Winnock et al., 2002) or a non significant trend to faster decline (Proust-Lima et al., 2015) depending on the cognitive tests under study.

Since the BIC of the models with two, three and four latent classes estimated on this ApoE sub-sample are very close, we also present the classification coming from the four latent class model for comparison with the analysis of the whole PAQUID dataset. The proportions of subjects classified in the four posterior latent classes are similar in the two samples (see captions of Web Tables 7 and 10), as well as the estimated class-specific mean Isaacs trajectories (part C of Web Figure 2). The main differences are a delayed intensity transition to Death in class 2 that leads to a much higher proportion of subjects that develop dementia before death and a lower transition intensity to Dementia in class 3 (the healthiest class). Besides, this model has a higher discriminatory ability, as shown in Web Table 10.

References

Winnock, M., Letenneur, L., Jacqmin-Gadda, H., Dallongeville, J., Amouyel, P. and Dartigues, J. F (2002). Longitudinal analysis of the effect of apolipoprotein E epsilon4 and education on cognitive performance in elderly subjects: the PAQUID study. *Journal of Neurology, Neurosurgery & Psychiatry* 72(6):794–7.
(a) Visits every 2 years

| Class Membership | \( \hat{\lambda}_1 \) | \( \hat{\lambda}_2 \) | \( \beta_1 \) | \( \beta_2 \) | ASE | ESE | Cover Rate |
|------------------|-----------------|-----------------|----------|----------|------|------|-----------|
| Baseline Risks of events | 3.20 | 3.20 | 0.01 | 0.2053 | 0.3377 | 0.95 |
| \( \lambda_{11} \) | 2.10 | 2.10 | 0.10 | 0.0009 | 0.0007 | 0.94 |
| \( \lambda_{12} \) | 3.50 | 3.50 | 0.5250 | 0.2848 | 0.94 |
| \( \lambda_{22} \) | 2.10 | 2.10 | 0.10 | 0.0009 | 0.0007 | 0.94 |
| \( \lambda_{21} \) | 3.50 | 3.50 | 0.5250 | 0.2848 | 0.94 |
| \( \lambda_{22} \) | 3.40 | 3.40 | 0.1666 | 0.1656 | 0.96 |
| \( \lambda_{21} \) | 0.11 | 0.11 | 0.0006 | 0.0006 | 0.95 |
| \( \lambda_{22} \) | 0.10 | 0.10 | 0.0005 | 0.0005 | 0.95 |
| \( \lambda_{21} \) | 3.10 | 3.10 | 0.5242 | 0.5292 | 0.93 |
| \( \lambda_{22} \) | 0.12 | 0.12 | 0.0089 | 0.0089 | 0.88 |
| \( \lambda_{21} \) | 0.11 | 0.11 | 0.0047 | 0.0047 | 0.89 |
| Event covariates | 0.02 | 0.02 | 0.1729 | 0.1708 | 0.94 |
| \( \gamma_1 \) | 0.07 | 0.07 | 0.1099 | 0.1117 | 0.94 |
| \( \gamma_2 \) | 0.47 | 0.47 | 0.2111 | 0.2301 | 0.92 |
| \( \gamma_3 \) | 0.00 | 0.00 | 0.3035 | 0.2133 | 0.96 |

(b) Visits every 4 years

| Class Membership | \( \hat{\lambda}_1 \) | \( \hat{\lambda}_2 \) | \( \beta_1 \) | \( \beta_2 \) | ASE | ESE | Cover Rate |
|------------------|-----------------|-----------------|----------|----------|------|------|-----------|
| Baseline Risks of events | 3.20 | 3.20 | 0.01 | 0.2053 | 0.3377 | 0.95 |
| \( \lambda_{11} \) | 2.10 | 2.10 | 0.10 | 0.0009 | 0.0007 | 0.94 |
| \( \lambda_{12} \) | 3.50 | 3.50 | 0.5250 | 0.2848 | 0.94 |
| \( \lambda_{22} \) | 2.10 | 2.10 | 0.10 | 0.0009 | 0.0007 | 0.94 |
| \( \lambda_{21} \) | 3.50 | 3.50 | 0.5250 | 0.2848 | 0.94 |
| \( \lambda_{22} \) | 3.40 | 3.40 | 0.1666 | 0.1656 | 0.96 |
| \( \lambda_{21} \) | 0.11 | 0.11 | 0.0006 | 0.0006 | 0.95 |
| \( \lambda_{22} \) | 0.10 | 0.10 | 0.0005 | 0.0005 | 0.95 |
| \( \lambda_{21} \) | 3.10 | 3.10 | 0.5242 | 0.5292 | 0.93 |
| \( \lambda_{22} \) | 0.12 | 0.12 | 0.0089 | 0.0089 | 0.88 |
| \( \lambda_{21} \) | 0.11 | 0.11 | 0.0047 | 0.0047 | 0.89 |
| Event covariates | 0.02 | 0.02 | 0.1729 | 0.1708 | 0.94 |
| \( \gamma_1 \) | 0.07 | 0.07 | 0.1099 | 0.1117 | 0.94 |
| \( \gamma_2 \) | 0.47 | 0.47 | 0.2111 | 0.2301 | 0.92 |
| \( \gamma_3 \) | 0.00 | 0.00 | 0.3035 | 0.2133 | 0.96 |

Web Table 1: Results of the simulation study comparing estimates of the two-latent-class joint linear markovian illness-death model for interval-censored data and the joint competing risks model based on 500 samples of 1000 subjects generated with a joint markovian illness-death model with visits every 2 years or every 4 years. ASE is the asymptotic standard error, ESE is the empirical standard error and the coverage rate is calculated from the 95% confidence interval.
Web Table 2: Results of the simulation study for the two-latent-class joint linear semi-Markovian illness-death model for interval-censored data based on 500 samples of 500 subjects generated with visits every 2 years or every 4 years. ASE is the asymptotic standard error, ESE is the empirical standard error and the coverage rate is calculated from the 95% confidence interval.
Web Table 3: Results of the simulation study of the three-latent-class joint linear markovian illness-death model for interval-censored data based on 500 samples of 500 subjects generated with visits every 2 years or every 4 years. ASE is the asymptotic standard error, ESE is the empirical standard error and the coverage rate is calculated from the 95% confidence interval.
(a) \( \lambda_{01}(t) = \lambda_{01}^0(t) \exp(\beta X_i + u_{01}) \)

(b) \( \lambda_{01}(t) = \lambda_{01}^0(t) \exp(\beta X_i + 0.2 u_{01}) \)

| Class Membership | \( \beta \) | \( \hat{\beta} \) | ASE | ESE | Cover Rate |
|------------------|---------|--------|-----|-----|----------|
| \( \zeta_1 \)    | 0.00    | -0.01  | 0.1370 | 0.1429 | 0.95     |
| \( \zeta_2 \)    | 0.00    | 0.00   | 0.1565 | 0.1497 | 0.95     |

| Baseline Risks of events | \( \lambda_{01}^0 \) | \( \hat{\lambda}_{01} \) | \( \beta \) | ASE | ESE | Cover Rate |
|--------------------------|--------------------|------|------|-----|-----|----------|
| \( \lambda_{101} \)     | 4.95               | 4.94 | 0.2855 | 0.2758 | 0.95     |
| \( \lambda_{102} \)     | 4.70               | 4.67 | 0.2490 | 0.2486 | 0.94     |
| \( \lambda_{103} \)     | 4.84               | 4.83 | 0.2356 | 0.2352 | 0.95     |
| \( \lambda_{201} \)     | 0.10               | 0.10 | 0.0003 | 0.0003 | 0.93     |
| \( \lambda_{202} \)     | 0.11               | 0.11 | 0.0003 | 0.0003 | 0.93     |
| \( \lambda_{203} \)     | 3.31               | 3.32 | 0.2617 | 0.2600 | 0.95     |
| \( \lambda_{211} \)     | 3.28               | 3.28 | 0.3222 | 0.3112 | 0.96     |
| \( \lambda_{212} \)     | 3.10               | 3.11 | 0.3212 | 0.3254 | 0.95     |
| \( \lambda_{213} \)     | 0.10               | 0.10 | 0.0007 | 0.0007 | 0.96     |
| \( \lambda_{221} \)     | 0.11               | 0.11 | 0.0008 | 0.0009 | 0.96     |
| \( \lambda_{222} \)     | 0.10               | 0.10 | 0.0009 | 0.0100 | 0.96     |
| \( \lambda_{223} \)     | 3.31               | 3.32 | 0.2617 | 0.2600 | 0.95     |

| Event covariates | \( \gamma_{01} \) | \( \hat{\gamma}_{01} \) | ASE | ESE | Cover Rate |
|------------------|------------------|---------|-----|-----|----------|
| \( \gamma_{01} \) | -1.07            | -1.06   | 0.1558 | 0.1617 | 0.92     |
| \( \gamma_{02} \) | -0.12            | -0.12   | 0.1386 | 0.1423 | 0.94     |
| \( \gamma_{12} \) | -0.03            | -0.04   | 0.1863 | 0.2032 | 0.93     |

| Latent process | \( \beta \) | \( \hat{\beta} \) | ASE | ESE | Cover Rate |
|----------------|---------|--------|-----|-----|----------|
| \( \beta_{01} \) | 32.05   | 32.08  | 0.2221 | 0.2190 | 0.96     |
| \( \beta_{02} \) | 30.99   | 31.03  | 0.2344 | 0.2320 | 0.96     |
| \( \beta_{03} \) | 28.97   | 29.02  | 0.2197 | 0.2215 | 0.96     |
| \( \beta_{11} \) | -3.08   | -3.05  | 0.1604 | 0.1602 | 0.96     |
| \( \beta_{12} \) | -5.60   | -5.56  | 0.1133 | 0.1080 | 0.96     |
| \( \beta_{13} \) | -7.69   | -7.66  | 0.1038 | 0.1071 | 0.94     |
| \( \beta_{21} \) | 4.83    | 4.83   | 0.1115 | 0.1191 | 0.94     |

| Cholesky transformation | \( \mathbf{U}(1,1) \) | \( \mathbf{U}(1,2) \) | \( \mathbf{U}(2,2) \) | \( \mathbf{U}(1,1) \) | \( \mathbf{U}(1,2) \) | \( \mathbf{U}(2,2) \) | \( \mathbf{U}(1,1) \) | \( \mathbf{U}(1,2) \) | \( \mathbf{U}(2,2) \) |
|-------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| \( \mathbf{U}(1,1) \)   | 2.24                | 2.22                | 0.1027              | 0.1027              | 0.93                |
| \( \mathbf{U}(1,2) \)   | -0.89               | -0.89               | 0.0581              | 0.0591              | 0.93                |
| \( \mathbf{U}(2,2) \)   | 0.45                | 0.45                | 0.0233              | 0.0243              | 0.93                |

| Measurement error | \( \sigma \) | \( \hat{\sigma} \) | ASE | ESE | Cover Rate |
|------------------|---------|--------|-----|-----|----------|
| \( \sigma \)    | 1.00    | 1.00   | 0.0137 | 0.0139 | 0.94     |

Models converged based on 500 samples

Web Table 4: Results of the simulation study of the misspecified three-latent-class joint linear markovian illness-death model for interval-censored data based on 500 samples of 500 subjects generated with visits every 2 years. The simulated transition intensity to dementia depends on (a) the individual random intercept or (b) the individual random slope.
| \( \lambda \) | \( \beta \) | \( \hat{\beta} \) | \( ASE \) | \( ESE \) | Cover Rate |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( \zeta_1 \) | 0.00 | -0.00 | 0.1158 | 0.1171 | 0.94 |
| \( \zeta_2 \) | 0.00 | 0.00 | 0.1385 | 0.1375 | 0.95 |
| \( \zeta_3 \) | 0.00 | 0.00 | 0.1488 | 0.1545 | 0.92 |
| \( \zeta_4 \) | 0.00 | 0.00 | 0.1962 | 0.1955 | 0.96 |
| \( \zeta_5 \) | 0.00 | 0.01 | 0.2021 | 0.2244 | 0.93 |

| \( \gamma \) | \( \beta \) | \( \hat{\beta} \) | \( ASE \) | \( ESE \) | Cover Rate |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( \gamma_01 \) | -1.07 | -1.08 | 0.0868 | 0.0878 | 0.93 |
| \( \gamma_02 \) | -0.12 | -0.13 | 0.0705 | 0.0731 | 0.93 |
| \( \gamma_{12} \) | -0.03 | -0.03 | 0.1106 | 0.1127 | 0.94 |

| \( \beta \) | \( \hat{\beta} \) | \( ASE \) | \( ESE \) | Cover Rate |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( \beta_{01} \) | 32.05 | 32.06 | 0.1643 | 0.1591 | 0.95 |
| \( \beta_{02} \) | 30.99 | 31.00 | 0.1901 | 0.1878 | 0.96 |
| \( \beta_{03} \) | 29.97 | 29.97 | 0.1904 | 0.2016 | 0.91 |
| \( \beta_{04} \) | 28.05 | 28.08 | 0.2321 | 0.2495 | 0.93 |
| \( \beta_{05} \) | 27.20 | 27.21 | 0.2290 | 0.2387 | 0.95 |
| \( \beta_{06} \) | 26.00 | 25.99 | 0.1721 | 0.1772 | 0.95 |
| \( \beta_{11} \) | -3.08 | -3.08 | 0.0774 | 0.0754 | 0.95 |
| \( \beta_{12} \) | -4.60 | -4.60 | 0.0948 | 0.0957 | 0.95 |
| \( \beta_{13} \) | -5.69 | -5.68 | 0.0882 | 0.0914 | 0.94 |
| \( \beta_{14} \) | -6.45 | -6.45 | 0.1139 | 0.1154 | 0.94 |
| \( \beta_{15} \) | -7.10 | -7.10 | 0.1127 | 0.1182 | 0.93 |
| \( \beta_{16} \) | -8.00 | -8.00 | 0.0805 | 0.0860 | 0.93 |
| \( \beta_X \) | 4.83 | 4.83 | 0.0861 | 0.0870 | 0.95 |

| \( \sigma_e \) | 1.00 | 1.00 | 0.0068 | 0.0068 | 0.95 |

*Models converged based on 456 samples*
Web Table 6: Estimation of the parameters of the four-latent-class joint change-point markovian illness-death model for semi-competing interval-censored events and longitudinal data (PAQUID data, N=3,525).

| Class Membership | \( \beta \) | SE  | CI inf | CI sup |
|------------------|-------------|-----|--------|--------|
| \( \pi_1 \)      | -0.40       | 0.23| -0.86  | 0.05   |
| \( \pi_2 \)      | 0.05        | 0.23| -0.40  | 0.50   |
| \( \pi_3 \)      | -1.32       | 0.23| -1.76  | -0.87  |

| Illness-death model | \( \gamma \) | SE  | CI inf | CI sup |
|---------------------|-------------|-----|--------|--------|
| Education \( \hat{\gamma}_{e1} \) | -1.18       | 0.12| -1.43  | -0.94  |
| Gender \( \hat{\gamma}_{g1} \)    | -0.07       | 0.11| -0.29  | 0.15   |
| Education \( \hat{\gamma}_{e2} \) | -0.29       | 0.07| -0.43  | -0.16  |
| Gender \( \hat{\gamma}_{g2} \)    | -0.70       | 0.06| -0.83  | -0.58  |
| Education \( \hat{\gamma}_{e12} \) | -0.00       | 0.05| -0.09  | 0.08   |
| Gender \( \hat{\gamma}_{g12} \)   | -0.54       | 0.08| -0.70  | -0.37  |

| Latent process | \( \tau \) | SE  | CI inf | CI sup |
|----------------|----------|-----|--------|--------|
| Class-specific | \( \tau_1 \) | 84.22| 0.86   | 82.52  | 85.91  |
| Change-point times | \( \tau_2 \) | 82.62| 0.44   | 81.75  | 83.48  |
|               | \( \tau_3 \) | 89.42| 1.12   | 87.23  | 91.61  |
|               | \( \tau_4 \) | 86.30| 0.62   | 85.08  | 87.52  |

| Education \( \beta_3 \) | 1.280 | 0.09 | 1.11 | 1.45 |
| Education \( \times t \ if \ t < \tau_g \) \( \beta_4 \) | -0.002 | 0.03 | -0.07 | 0.07 |
| Education \( \times t \ if \ t > \tau_g \) \( \beta_5 \) | -0.004 | 0.04 | -0.09 | 0.08 |
| Gender \( \beta_6 \) | 0.090 | 0.05 | 0.00 | 0.18 |

Web Table 7: Mean probabilities (in percentages) to belong to each class according to the posterior classification, allocating 392 (11.12%) subjects in class 1, 1160 (32.91%) to class 2, 343 (9.73%) to class 3 and 1630 (46.24%) to class 4 (PAQUID data, N=3,525).

| Class | 1     | 2     | 3     | 4     |
|-------|-------|-------|-------|-------|
| 1     | 71.36 | 21.92 | 0.18  | 6.54  |
| 2     | 12.83 | 61.54 | 1.24  | 24.39 |
| 3     | 0.01  | 0.47  | 79.11 | 20.40 |
| 4     | 0.96  | 19.11 | 12.96 | 66.97 |
Web Table 8: Class-specific description of the posterior classes of the four-latent-class joint change-point markovian illness-death model, for gender and educational level, in percentages (PAQUID data, N=3,525).

| Class | CEP=0 | CEP=1 | men | women |
|-------|-------|-------|-----|-------|
| 1     | 23.2  | 76.8  | 40.1| 59.9  |
| 2     | 29.7  | 70.3  | 40.7| 59.3  |
| 3     | 42.9  | 57.1  | 42.0| 58.0  |
| 4     | 38.0  | 62.0  | 43.8| 56.2  |

Web Table 9: Estimation of the parameters of the two-latent-class joint change-point markovian illness-death model, adjusted on gender, educational level and ApoE4 (sub-sample of PAQUID cohort with ApoE4 measurement, N=619).
Web Table 10: Mean probabilities (in percentages) to belong to each class according to the posterior classification, allocating 103 (16.64%) subjects in class 1, 145 (23.42%) to class 2, 64 (10.34%) to class 3 and 307 (49.60%) to class 4 (sub-sample of PAQUID cohort with ApoE4 measurement, N=619).

| Class | 1    | 2    | 3    | 4    |
|-------|------|------|------|------|
| 1     | 76.21| 9.30 | 0.35 | 14.14|
| 2     | 4.63 | 77.36| 1.99 | 16.02|
| 3     | 0.00 | 0.00 | 87.70| 12.29|
| 4     | 6.42 | 8.88 | 10.37| 74.33|
A) Predicted means of Isaacs scores given the random effects and the classes by age range (x) and weighted means of the observed scores (solid lines) with their 95% confidence intervals (dashed lines) for each class.

B) Class-specific predicted cumulative incidences estimated by the joint latent class illness-death model (thick lines) and by the weighted semi-parametric illness-death model (thin lines), for each class (class 1: dashed line, class 2: dotted line, class 3: dotdashed line, class 4: solid line).

Web Figure 1: Goodness-of-fit assessment of the four-latent-class joint change-point illness-death model estimated on PAQUID data (N= 3,525).
A) Class-specific transition intensities of the illness-death model for men, ApoE4 non-carrier, with a low level of education.

B) Class-specific cumulative incidences of the illness-death model for men, ApoE4 non-carrier, with a low level of education.

C) Class-specific Isaacs trajectories for men, ApoE4 non-carrier, with a low level of education.

Web Figure 2: Class-specific estimated transition intensities, cumulative incidences and mean longitudinal trajectories of the joint change-point latent class illness-death mixed model for each class (class 1: dashed line, class 2: dotted line, class 3: dotdashed line, class 4: solid line, sub-sample from PAQUID cohort with ApoE4 measurement, N= 619).