Department of Biostatistics
Erasmus University Medical Center

Consultancy Report

Title	Joint modeling for kidney transplant patient monitoring
Version	1
Date	12 January 2017
Author	Anirudh Tomer, Dimitris Rizopoulos, Ewout Steyerberg
Collaborators	Jesper Kers, Hessel Peters Sengers

Visiting address: Wytemaweg 80, 3015 CN Rotterdam, Na building (room Na-2418)
Postal address: PO Box 2040, 3000 CA Rotterdam, the Netherlands
Web: http://www.erasmusmc.nl/biostatistiek/

1 Introduction

In this report we present the analysis of the kidney transplant data set of Academic Medical Center, University of Amsterdam. The data set contains information about 239 patients who received kidney transplants. Each patient was uniquely identified by their study number('amctx'), which also had a one to one correspondence with patient number ('zis'). While some information such as donor gender, receiver gender, blood pressure etc. were measured only once, serum creatinine and urinary protein-creatinine ratio(pcr) were measured repeatedly over time till the patient either had a transplant failure (death of patient/graft failure), or was not followed up anymore. The analysis presented in this report was motivated by the following questions:

1. How does the slope of serum creatinine over time predict death-censored graft failure?
2. Is there an added value of the slope of urinary pcr over time to predict death-censored graft failure?

Since both serum creatinine and urinary pcr are internal time dependent covariates, to answer the aforementioned questions, we used a joint modeling approach [Rizopoulos, 2012] rather than a time dependent Cox model approach. We used the R package JMbayes (R version 3.3.1) for estimating the parameters in the joint model.

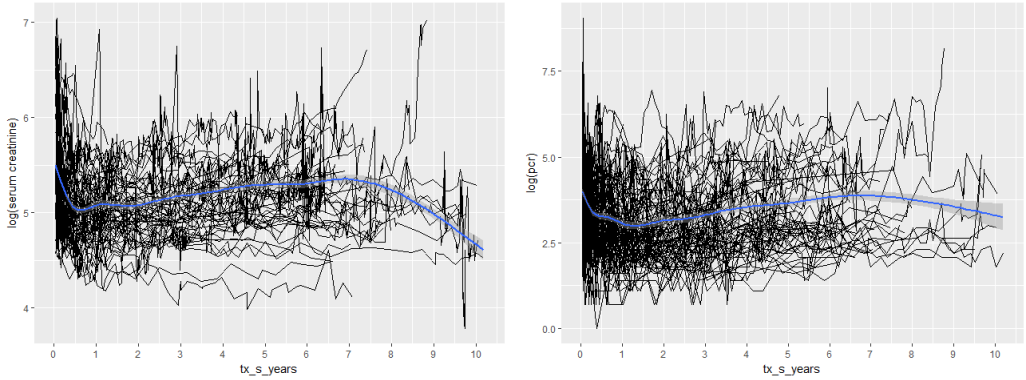
2 Multivariate longitudinal analysis of serum creatinine and urinary pcr

We first created a multivariate longitudinal model for serum creatinine and urinary pcr to check which covariates were useful in predicting the two longitudinal outcomes. Serum creatinine measurements were available for 239 patients whereas urinary pcr measurements were available for only 238 patients. For multivariate longitudinal modeling of these two measurements we required equal number of subjects and hence subject with study number 346 was not considered in the analysis. Secondly for certain patients, there were multiple unequal creatinine measurements at the same time point. Based on the suggestions from AMC, we kept only the first of such measurements. For using the covariate receiver age, we used its value at the first follow up rather than using it as a time varying covariate. This was done because we had another time varying covariate in the model called 'tx_s_days' (days from transplantation until the measurement). We modeled non-linear evolutions of log (creatinine) and log (pcr) over time (years from transplantation until the measurement) while controlling for the other covariates present in the data set. For the non linear evolutions we used splines with knots based on graphical analysis of the measurements. The choice of log transforming both of the longitudinal responses was based on graphical analysis of residuals as well as of the trend plots of the response (Figure 1). We will now present the covariates with which we fitted the model for the two longitudinal responses. For brevity, the parameter estimates are discussed only for the joint model (Section 4).

2.1 Covariates for serum creatinine

The sub-model we used for serum creatinine was an additive model consisting of the following covariates:

- rec_age: Receiver age at the first follow up.
- rec_gender: Receiver gender.
- d_age: Donor age.
- tx_dgf: Delayed graft function after transplantation.
- is_cni: Use of a calcineurin inhibitor.
- d_bmi: Donor BMI.
- rec_bmi: Receiver BMI.



(a) Observed evolutions log(creatinine). (b) Observed evolutions log(pcr).

Figure 1: Observed evolutions of log(creatinine) and log(pcr) for a randomly selected set of patients. X axis is years from transplantation until the measurement.

- tx_hla: Number of HLA A, B and DR mismatches between donor and recipient.
- tx_previoustx: Previous transplantation before the current transplantation.
- tx_pra: Panel reactive antibody percentage before transplantation.
- tx_cit: Cold ischemia time (minutes).
- tx_dial_days: Number of days on dialysis prior to transplantation.
- tx_dm: Diabetes mellitus
- tx_s_days: Natural cubic spline with knots at 100, 300 and 1000 days. In the model we scaled days to years for computational reasons.

The model also had a random intercept and a spline random effect part with knots at 100 and 300 days.

2.2 Model for urinary pcr

The model we used for urinary pcr was an additive model consisting of the following covariates:

- ah_ace: Use of ACE inhibition.
- ah_arb: Use of angiotensin-2 receptor blockers.
- ah_raasi: Use of renin-angiotensin-aldosterone inhibitors.
- rec_age: Receiver age at the first follow up.
- d_age: Donor age.
- d_bmi: Donor BMI.
- rec_bmi: Receiver BMI.
- d_type: Donor type.
- tx_previoustx: Previous transplantation before the current transplantation.
- tx_pra: Panel reactive antibody percentage before transplantation.
- tx_hla: Number of HLA A, B and DR mismatches between donor and recipient.
- tx_cit: Cold ischemia time (minutes).
- tx_dial_days: Number of days on dialysis prior to transplantation.

- tx_dm: Diabetes mellitus
- ah_nr: Number of anti-hypertensive medicaments.
- tx_s_days: Natural cubic spline with knots at 100, 200 and 350 days. In the model we scaled days to years for computational reasons.

The model also had a random intercept and a spline random effect part with knots at 100 and 200 days.

3 Time to event analysis of death-censored graft failure

We used a Cox model for time to event analysis of death-censored graft failure. Out of the 238 subjects considered for analysis, only 44 subjects had a failures and rest were censored. Given the large number of censored observations we had a low statistical power to detect significant effects of covariates in a large extensive model. Based on the suggestion of the team at AMC and supplementing it with our own analysis we took the additive effect of the following covariates in our model:

- d_age: Donor age.
- tx_previoustx: Previous transplantation before the current transplantation.
- rec_bmi: Receiver BMI.
- d_type: Donor type.
- ah_diur: Use of diuretics.
- tx_dgf: Delayed graft function after transplantation
- tx_hla: Number of HLA A, B and DR mismatches between donor and recipient.
- rec_age: Receiver age at the first follow up.
- tx_pra: Panel reactive antibody percentage before transplantation.

4 Joint model for time to event analysis of death-censored graft failure

Using the longitudinal and survival submodels described above we fitted a joint model. In the joint model the association between the survival and longitudinal outcomes was considered via the value of log(creatinine), log(pcr) and the slopes of each of these outcomes at any given point in time. Since we had very few subjects who had events, to estimate the parameters in the survival submodel we used a Bayesian ridge approach. Table 1 shows the parameter estimates for the survival submodel in this joint model. It can be seen that among the baseline covariates 'ah_diur', 'tx_dgf', 'tx_hla', 'tx_pra', 'd_age' and 'rec_age' are not significant in them model. However, since they have clinical relevance we decided to keep them in the model anyway. Secondly neither log(pcr) nor its slope was found to be significant in the model. The association of slope of log(creatinine) was also not found to be significant. Using a backward elimination approach we first removed slope of log(pcr) and then refitted the model. We found associations with log(pcr) and slope of log(creatinine) to be still insignificant. Further we fitted two more models, one having associations of event time outcome with log(creatinine) and log(pcr), and another having associations of event time outcome with log(creatinine) and slope of log(creatinine). In both of these models only the association with log(creatinine) was significant.

While the obvious choice for a final model was: One longitudinal outcome log(creatinine) and its association with time to event, we decided to keep the association of slope of log(creatinine) as well, because the primary research question was "How does the slope of serum creatinine over time predict death censored graft failure?". The corresponding model's parameter estimates for

	Mean	Std.Dev	2.5%	97.5%	P
d_age	0.001	0.015	-0.028	0.032	0.946
tx_previoustx: Yes	1.157	0.432	0.252	2.006	0.002
rec_bmi	0.107	0.037	0.030	0.182	0.016
d_type: LRD	-0.221	0.365	-1.149	0.341	0.536
d_type: LURD	0.158	0.354	-0.480	1.028	0.634
d_type: NHBD	-1.265	0.509	-2.265	-0.194	0.014
ah_diur: Yes	-0.055	0.269	-0.636	0.455	0.830
tx_dgf: Yes	0.041	0.271	-0.503	0.638	0.852
tx_hla: 1	0.062	0.399	-0.718	0.948	0.866
tx_hla: 2	0.033	0.338	-0.721	0.772	0.912
tx_hla: 3	0.198	0.388	-0.515	1.113	0.598
tx_hla: 4	-0.112	0.349	-0.835	0.569	0.682
tx_hla: 5	-0.113	0.362	-0.954	0.571	0.742
tx_hla: 6	-0.017	0.290	-0.651	0.558	0.984
rec_age	0.016	0.014	-0.013	0.044	0.216
tx_pra	0.004	0.009	-0.014	0.021	0.640
log(pcr)	0.210	0.140	-0.038	0.489	0.130
Slope (log(pcr))	-0.029	0.125	-0.270	0.220	0.814
log(creatinine)	1.691	0.395	0.843	2.397	0.004
Slope (log(creatinine))	0.448	0.391	-0.097	1.311	0.212

Table 1: Parameter estimates for the survival submodel in the joint model with associations for both log(creatinine), log(pcr) and their slopes.

	Mean	Std.Dev	2.5%	97.5%	P
Intercept	4.747	0.225	4.308	5.178	0.000
rec_age (at first follow up)	-0.004	0.002	-0.008	-0.001	0.010
rec_gender: Male	0.209	0.040	0.128	0.288	0.000
d_age	0.007	0.002	0.004	0.010	0.000
tx_dgf: Yes	0.133	0.049	0.038	0.236	0.004
is_cni: Yes	0.216	0.108	-0.006	0.418	0.056
d_bmi	-0.003	0.005	-0.012	0.006	0.538
tx_hla: 1	0.037	0.087	-0.135	0.197	0.666
tx_hla: 2	0.030	0.074	-0.114	0.172	0.684
tx_hla: 3	0.041	0.067	-0.088	0.173	0.528
tx_hla: 4	0.022	0.063	-0.103	0.144	0.718
tx_hla: 5	-0.010	0.059	-0.124	0.11	0.864
tx_hla: 6	-0.103	0.046	-0.191	-0.008	0.038
tx_previoustx: Yes	0.010	0.059	-0.103	0.125	0.862
tx_pra	0.002	0.001	-0.001	0.005	0.218
tx_cit	1.561e-04	4.553e-05	6.621e-05	2.441e-04	0.002
tx_dial_days	1.229e-05	1.625e-05	-1.747e-05	4.459e-05	0.466
tx_dm: Yes	-0.119	0.059	-0.239	-2.938e-04	0.05
rec_bmi	0.003	0.005	-0.006	0.013	0.506
Spline: 1	-0.243	0.035	-0.314	-0.174	0.000
Spline: 2	0.021	0.072	-0.117	0.159	0.764
Spline: 3	-0.448	0.192	-0.828	-0.093	0.020
Spline: 4	0.174	0.332	-0.454	0.808	0.610

Table 2: Parameter estimates for the longitudinal submodel. The outcome here is log(creatinine). The joint model has associations for both log(creatinine) its slope.

longitudinal and survival submodel for are shown in Table 2 and Table 3 respectively.

As we can see in Table 2, receiver age, receiver gender, donor age, need for dialysis within first week after transplant ('tx_dgf'), use of calcineurin inhibitor ('is_cni'), cold ischemia time ('tx_cit'), diabetes mellitus ('tx_dm') are all strongly related with log (serum creatinine). The interpretation for 'tx_dm' is that at any given time point, a patient with diabetes who takes OAM or insulin will have log(creatinine) level lower by 0.119 units if he/she weren't taking OAM or insulin. Interpretation for other categorical and continuous covariates can be done in the same way because the model has only the additive effects of the aforementioned covariates. The fitted evolution of log(creatinine) for a hypothetical kidney transplant recipient can be seen in Figure 2. This hypothetical patient has covariates levels equal to the median values of the covariates for patients in the data set at hand.

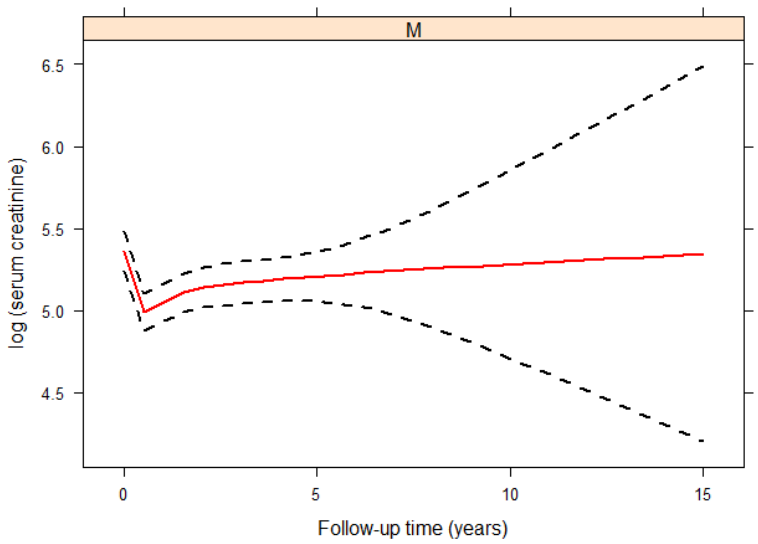


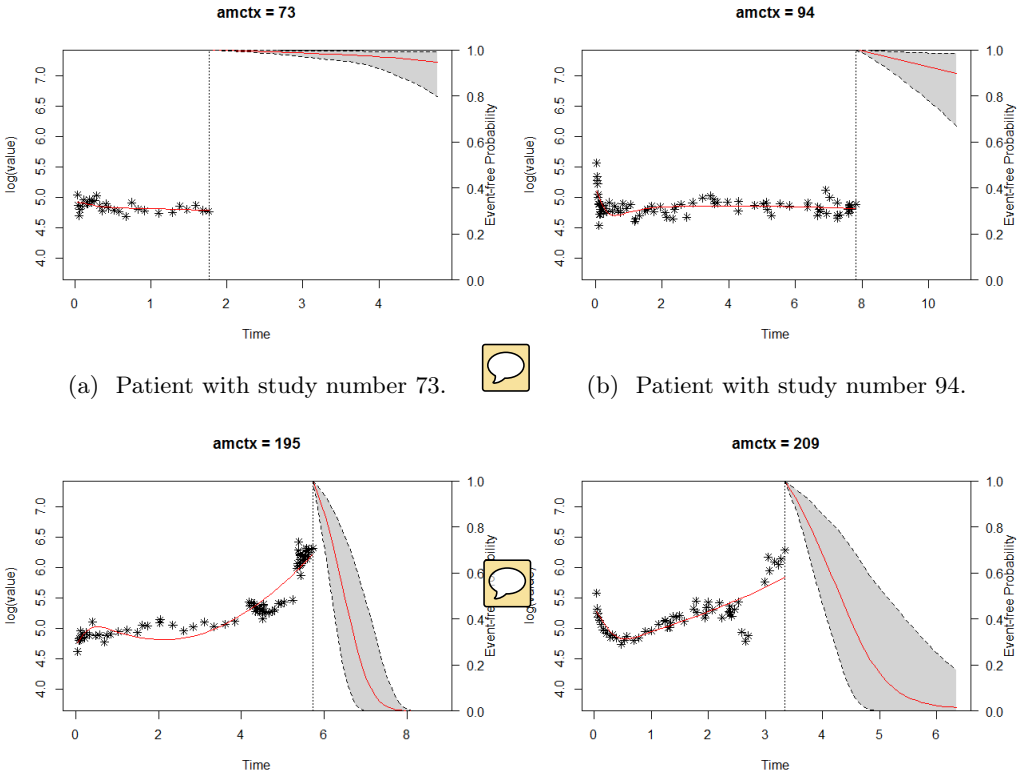
Figure 2: Fitted longitudinal profile of a patient with rec_gender = Male, rec_age = 53 years, d_age=52 years, d_bmi = 24, rec_bmi = 25, tx_dial_days = 1100, tx_cit = 960, tx_pra = 0, tx_dm = No, tx_previoustx = No, tx_hla = 6, is_cni = Yes, tx_dgf = No

In Table 3 we can see the effect of log(creatinine) and its slope on time to failure. To interpret the associations in the survival submodel, let us imagine a patient whose serum creatinine levels at some time point are 150 umol/L. If the patient instead had twice of these serum creatinine levels at that time point, i.e. 300 umol/L, then his hazard of failure had been 4.3 times more (95% CI [3.5, 5.3]). Although we also have association with slope of log(creatinine) in the model as well, it is not significant. Thus for all practical purposes the aforementioned interpretation of association of log(serum creatinine) remains valid.

To understand the impact of log(creatinine) levels on survival of a patient, we present in Figure 3a to Figure 3d the observed and fitted longitudinal profiles for log (creatinine) of 4 patients, and their survival probabilities 3 years after their last follow up time. As we can see patient with study number 73 and 94 have more or less a flat trajectory of log(creatinine) and consequently their survival probabilities do not decrease rapidly as compared to patients 195 and 209. Figure 4a to Figure 4d show the predicted longitudinal profiles for the 4 patients 3 years after the loss of follow up. As expected based on the observed profiles of the latter 2 patients, their log(creatinine) levels are predicted to consistently increase over the next 3 years.

	Mean	Std.Dev	2.5%	97.5%	P
d_age	-0.004	0.005	-0.015	0.009	0.394
tx_previoustx: Yes	1.175	0.162	0.816	1.513	0.000
rec_bmi	0.102	0.014	0.070	0.133	0.000
d_type: LRD	-0.286	0.188	-0.576	0.085	0.158
d_type: LURD	0.062	0.134	-0.240	0.365	0.504
d_type: NHBD	-1.251	0.189	-1.566	-0.776	0.000
ah_diur: Yes	-0.204	0.120	-0.389	0.070	0.146
tx_dgf: Yes	0.073	0.111	-0.168	0.286	0.460
tx_hla: 1	0.104	0.159	-0.228	0.435	0.488
tx_hla: 2	-0.007	0.128	-0.305	0.231	0.822
tx_hla: 3	0.282	0.195	-0.110	0.588	0.198
tx_hla: 4	-0.166	0.155	-0.483	0.127	0.310
tx_hla: 5	-0.186	0.173	-0.554	0.166	0.270
tx_hla: 6	-0.046	0.118	-0.238	0.220	0.662
rec_age	0.019	0.005	0.008	0.030	0.008
tx_pra	0.003	0.003	-0.003	0.010	0.268
log(creatinine)	2.115	0.146	1.801	2.399	0.000
Slope (log(creatinine))	0.482	0.254	-0.016	1.034	0.058

Table 3: Parameter estimates for the survival submodel in the joint model with associations for both log(creatinine) its slope.



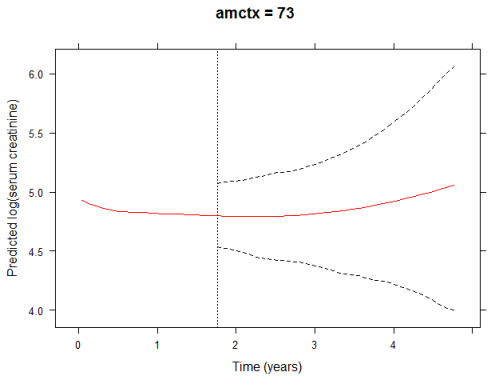
(a) Patient with study number 73.

(b) Patient with study number 94.

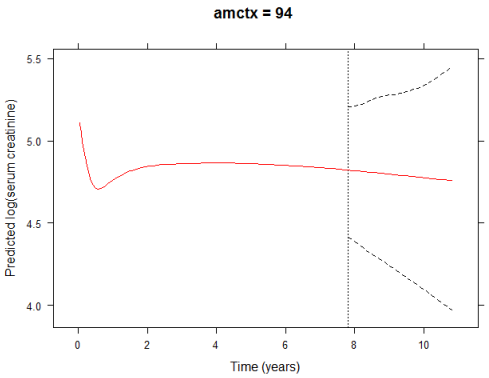
(c) Patient with study number 195.

(d) Patient with study number 209.

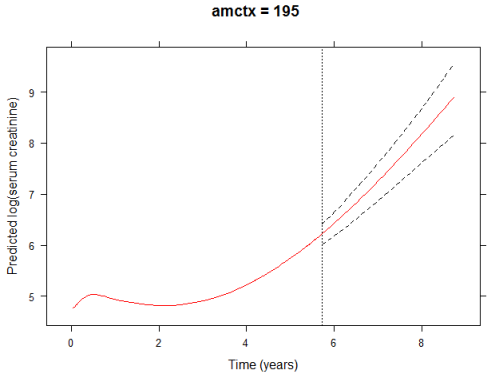
Figure 3: Dynamic predictions of survival probabilities up to 3 years after loss of follow up. Time is in years, and the trend plot show log(creatinine).



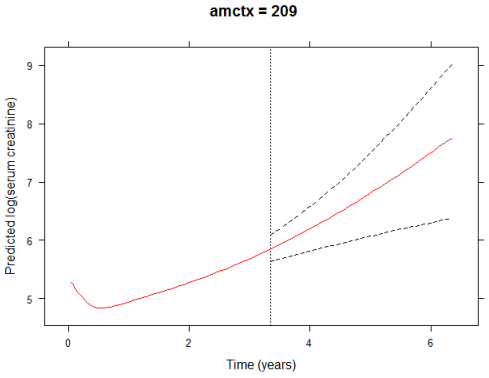
(a) Patient with study number 73.



(b) Patient with study number 94.



(c) Patient with study number 195.



(d) Patient with study number 209.

Figure 4: Dynamic predictions of log(creatinine) up to 3 years after loss of follow up.

References

[Rizopoulos, 2012] Rizopoulos, D. (2012). *Joint models for longitudinal and time-to-event data: With applications in R.* CRC Press.