

A personalized screening strategy to monitor the development of chronic allograft failure in renal transplant recipients

Hessel Peters-Sengers^{1,2,*}, Anirudh Tomer^{3,*}, Sandrine Florquin⁴, Ewout W. Steyerberg^{5,6}, Frederike J. Bemelman¹, Dimitris Rizopoulos^{3,\$}, Jesper Kers^{4,7,\$}

¹Division of Internal Medicine, Renal Transplant Unit, Academic Medical Center, Amsterdam, The Netherlands

²Center for Experimental and Molecular Medicine (CEMM), Academic Medical Center, Amsterdam, The Netherlands

³Department of Biostatistics, Erasmus University Medical Center, Rotterdam, The Netherlands

⁴Department of Pathology, Academic Medical Center, Amsterdam, The Netherlands

⁵Department of Public Health, Erasmus University Medical Center, Rotterdam, The Netherlands


⁶Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

⁷Van 't Hoff Institute for Molecular Sciences (HIMS), University of Amsterdam, Amsterdam, The Netherlands

*,\$ Authors contributed equally to the study

Introduction

Renal transplantation is the preferred modality of renal replacement therapy available to date. Compared to dialysis, renal transplantation has a favourable risk profile for (cardiovascular) outcomes and the quality of life for recipients of a transplant is better than dialyzing patients.¹⁻³ Short-term graft survival is excellent⁴, but unfortunately, long-term graft survival has not benefitted from recent treatment regimens to the same extent⁵.

In an effort to stratify patients who are at increased or decreased risk for graft failure, a multitude of studies investigated the value of molecular biomarkers and clinical algorithms for the prediction of long-term graft failure. In clinical practice however, only very few of these markers are being used on a routine basis. The best example of markers that are practically  in every transplant outpatient clinic are serum creatinine as a readout for the glomerular filtration rate and urinalysis to determine protein leakage as a readout for glomerular and/or tubular dysfunction.^{6,7} Although these markers are determined on each outpatient visit, hardly any study determined the dynamic temporal association between serum creatinine and urinary protein content trajectories with the development of graft failure. In fact, the majority of studies, including randomized-controlled trials in kidney transplantation, model renal function under the paradigm that renal function declines in a linear pattern over time.⁸ A recent study by Ferro and colleagues highlighted that the contrary was true, namely that 87% of their included patients showed nonlinearity or nonprogression.⁹ Patients with subsequent graft failure were more likely to have episodes of rapid progression and less likely to have episodes of nonprogression. Similar dynamical data are not known for urinalysis, but we might assume that similar patterns of rapid progression and nonprogression here associate with outcome as well. These data corroborate on the fact that in renal transplantation, patients have to be continuously monitored and therefore prediction models, contrary to static models, have to be designed as such.

An interesting statistical model to simultaneously investigate longitudinal measurement of renal function and urinalysis and time-to-event graft survival data is the joint model. Joint models typically combine mixed effects models for repeated measures and Cox models for survival data. Recently, Rizopoulos and colleagues described an extension to the joint model that can be applied to monitoring of renal transplant recipients as well.¹⁰ This has been proven in a cohort of patients that had undergone aortic valve allograft replacement and were followed-up with echocardiographic measurements of aortic valve gradients over time (repeated measures submodel) to predict the composite outcome of re-operation and death (survival submodel) with a joint model.¹¹ They then used this joint model to calculate personalized screening intervals depending on the patients' longitudinal aortic valve gradient trajectories. In a simulation study, they compared how the personalized screening intervals performed as compared to the fixed-planning screening intervals and they could show a reduction in the number of screening time-points and a reduction in the absolute error for the optimal intervention time-point when using the personalized screening strategy compared to a predefined "one

size fits all" screening approach. The reduction in screening time-points when using the personalized screening strategy can be directly translated to a reduction in screening costs per individual.¹¹

We had two aims in the current study. Firstly, we wanted to create a joint model that included static baseline clinical data and dynamic longitudinal trajectories of serum creatinine and urinary protein-creatinine ratios to predict death-censored graft failure. Secondly, we wanted to use the joint model to test whether we could construct a personalized monitoring strategy and compare it to the fixed-term "one size fits all" monitoring strategy that is currently in use.

Methods

To assess the applicability of a personalized monitoring strategy, we performed a single-center retrospective cohort study in the Academic Medical Center (AMC), a tertiary referral hospital in Amsterdam, the Netherlands. The electronic patient database was used to collect all relevant data. All information was processed anonymously.

Study population

We screened the records of 239 end-stage renal disease patients that underwent renal transplantation at our institute from 200? to 20... The inclusion criteria for the study were: age at baseline ≥ 18 years who had >1 additional serum creatinine (SCr, $\mu\text{mol/L}$) and 24-hour urine collection to calculate protein:creatinine ratio (PCR, g/mol) during follow-up. SCr measurements were available for 239 patients whereas PCR measurements for 238 patients, leaving 238 subjects for the multivariate longitudinal model. All subjects were followed until Month 20...? Initial immunosuppressive therapy consisted of steroids combined with mycophenolate mofetil or mycophenolic acid and a calcineurin inhibitor, mostly tacrolimus but also cyclosporine. Alternatively, a combination of steroids, tacrolimus, and sirolimus was used. Donor kidneys were acquired through allocation by the Eurotransplant allocation program, Leiden, The Netherlands.¹²

Comment [HPS1]: Jesper?

Comment [HPS2]: Jesper?

Measures and outcomes

Included in the database were potential predictors for kidney function and graft failure. Extracted were: donor age, donor gender, donor body mass index (BMI), donor type (living; brain death (DBD); controlled circulatory death (DCD)); number of human leukocyte antigen (HLA) A, B, DR mismatches, cold ischemic time, panel reactive antibodies (PRA) before transplantation, immunosuppressive regimens at 3 months after transplantation, recipient dialysis vintage, recipient blood pressure, recipient age, recipient sex, recipient BMI, use of medication at 3 months after transplantation (anti-hypertensives; diuretics; ACE/ACEi/ARB inhibitors; beta blockers; calcium channel blockers; statins; antidiabetic medication or insulin. These were measured only once, whereas SCr and PCR were measured repeatedly over time till the patient either had a transplant failure (death or graft failure), or was not followed up anymore. If there were multiple SCr measurements per day, we kept the first of such measurements. We evaluated death-censored graft survival, defined as graft loss leading to

dialysis treatment as event and censored for death with a functioning graft. Additionally, we used SCr and PCR to study the longitudinal markers for kidney function. Delayed graft function was defined as need for dialysis treatment within seven days after transplantation.

Data analytic strategy

For comparing the fixed schedule with personalized schedule, we first fitted the joint model (JM). ~~The joint model referred to a submodel for death-censored graft failure and a submodel for longitudinal data of SCr and PCR.~~ A mixed model was applied for the longitudinal submodel of SCr and PCR, tested for assumption of nonlinearity and homoscedasticity. Assumption of homoscedasticity was violated, tested by graphical analysis of residuals as well as the fit of trend plots of the evolution of SCr and PCR, and solved by log-transformation. Due to non-linearity of evolution of SCr and PCR for all transplant patients, we included b-splines for time after transplantation with three internal knots at 30, 80, and 365 days, and two boundary knots at 14 days and 6 years. Random intercept and a random effect for slope of SCr and PCR was included. Other covariates to be included in the longitudinal submodel were based on the literature and clinical expertise. The quantitative transplantation characteristics are standardized to avoid convergence issues in parameter estimation.¹⁰ For the Cox cause-specific submodel of death-censored graft failure, proportionality and linearity were checked. Given the relatively low number of graft failure events to build a large extensive survival model, we chose to include covariates based on clinical expertise to reduce type II error. To further stabilize the regression coefficients in the Cox submodel, a Bayesian ridge approach was used.

Comment [HPS3]: Is this right?

Comment [HPS4]: Is this right
Anirudh?

Using the multivariate longitudinal and survival submodels as described above, we fitted the JM linked through the common random effects using the Bayesian methodology (see supplementary).^{10,13} Included in the JM were the impact of log(SCr) value and log(SCr) velocity, log(PCR) value and log(PCR) velocity, and transplantation characteristics to estimate the hazard of death-censored graft failure. While keeping all static transplant covariates in the model, a backward elimination approach was used to check which of the longitudinal markers remained significant in the joint model. We also calculated time dependent area under the receiver operating curve (tAUC) periodically at an interval of 6 months.^{14,15} Based on the predictive ability of SCr and PCR, the final JM was chosen to compare the fixed screening schedule with personalized schedule.¹¹ From this JM, we obtained the full specification of the joint distribution of SCr (value and velocity) and time to death-censored graft failure.

For estimation of the personalized schedules, the JM was used to define a patient-specific posterior predictive distribution of time of graft failure, given the observed SCr measurements. The optimal time of the next SCr measurement is the one at which the expected information gained from an extra SCr measurement is maximum. We considered the fixed schedule for measurement of SCr levels, common for all patients, to be 20 times in the first year and every three months thereafter. In order to create reasonable predictions, SCr measurements for the first 3 months are taken as per the fixed schedule. Since the SCr measurements are already taken for the kidney transplant patients, we conduct a small simulation in order to demonstrate the efficacy of the personalized schedules. From

the transplant dataset we sampled 625 patients, which are further split into a training (575 patients) and test (50 patients) part. From the training dataset, we obtained the posterior distribution of the parameters of the JM. Then, iteratively, we scheduled SCr measurements for the test dataset, until the dynamic risk of graft failure became larger than the threshold.¹⁶ The threshold dictates the amount of 'intervention time' at hand between intervention and graft failure. In this simulation we evaluated two thresholds, namely 5% and 2.5% risk of graft loss in 6 months. The difference between the time of intervention due to the schedule (fixed vs personalized) and the true intervention time was denoted as intervention offset. In addition, the time at hand between the observed intervention time and the time of graft failure was denoted as the failure offset. Ideally, both intervention and graft failure offsets should be zero (see supplemental I for a detailed description of the personalized schedule methods).

Continuous variables are presented as mean with standard deviation (SD). Kaplan Meier was used to estimate death-censored graft survival. Results from Cox regression are presented as hazard ratios (HRs) with 95% confidence intervals (CIs). Significance levels were set at the 5% level. Analyses were conducted using R (version 3.4.2)¹⁷ with the JMbays package (version 0.8-70)¹⁴ and survival package (version 2.41)¹⁸.

Results

Sample characteristics

Table 1 shows baseline characteristics of 239 kidney transplant patients and donors. Majority were recipients of deceased donors (74.1%). Mean recipient age was 50.7 (SD 12.7) years, majority firstly transplanted (84.5%), and comorbidities (diabetes 15.5%). In follow-up period we included a total of 13189 SCr measurements and 9616 PCR measurements. The median number of repeated SCr and PCR measurements per patient were 45 and 37, respectively. For SCr, 95% of the observations were taken before 6 years, while for PCR they were taken before 5.4 years. The median time between two SCr measurements was 10 days, while for PCR 14 days.

Delayed graft function rate was 32.2%, and higher for deceased donor transplants compared to living donor (42.9% vs 1.6%, $p < 0.001$). Figure 1 illustrates death-censored graft survival. At one year, 97.9% still had a functioning graft (95%CI 96.1-99.7), and this was 83.9% (95%CI 78.2-89.6) at 5 years.

Results Joint Model

Out of 239 patients, we use the data of only those 238 patients for whom both PCR and SCr data is available. Table 2 and Table 3 summarizes the regression coefficients of the longitudinal submodel for SCr and PCR, respectively. Since the quantitative variables are standardized, the effect sizes correspond to one standard deviation increase. Significantly associated with evolution of SCr were donor age, donor type, recipient age, recipient gender, recipient diabetes, recipient anti-hypertensive medicaments, and delayed graft function. It were only donor age and recipient BMI that were

Comment [HPS5]: We can also subdivide this into 1) longitudinal submodel, 2) graft failure submodel etc

significantly associated with evolution of PCR. For interpretation, Figure 2 and 3 shows the fitted evolution of SCr and PCR, respectively, from time of transplantation according to a female recipient of 50.7 years old of a first living donor kidney aged 49.7 years, and at the median of other variables. Table 4 summarizes the hazard ratios for the submodel of death-censored graft failure. The SCr levels were strongly associated with the hazard of graft failure: for a given patient at any time point, if the SCr levels increased with 25% and other variables remained the same, the hazard ratio of graft failure increased 1.43 times (adjusted HR 1.43, 95%CI 1.27-1.59, $p < .001$). Slope of SCr did not reach significance, interpreted as for patients having the same value for SCr and keeping other variables constant, the hazard for a 25% increase in the current slope of the evolution of SCr did not significantly increased the hazard for graft failure (adjusted HR 1.05 95%CI 0.99-1.10, $p = 0.082$). PCR (both value and velocity) were not significantly associated with graft failure. Including the longitudinal SCr and PCR, previous transplantation, HLA mismatches, CIT, and dialysis vintage were also not significantly associated. Figure 3 shows dynamic predictions of both graft survival and evolution of SCr of 2 randomly chosen transplant patients. To further verify if PCR were required in the model in presence of both log SCr levels and velocity, we fitted two more JMs. In the first JM, only SCr measurements (value and velocity) were included, and in the second JM only PCR measurements (value and velocity). Table 5 and Figure 4 summarizes the predictive ability with tAUCs of the longitudinal markers of the different JMs. Both SCr and PCR in the JM performed the same as the model with only SCr to discriminate between patients who obtained graft failure events. Therefore, the JM including only SCr measurements was chosen for personalized screening analyses, including SCr velocity although not significant. Except for year 3, all the tAUC of SCr (including value and velocity) were > 0.8 , indicating good discrimination.

Comment [HPS6]: I included the interpretation although not significant because it is still relevant.

1 comment: is there also a timeline to this slope? Like 25% increase within year?

Comment [HPS7]: Anirudh: Maybe perform an additional Wald test to check the overall effect of the marker? See 5.1.3 page 105/106 in book Dimitris.

Results Personalized vs fixed schedules

Figure 5, Figure 6, and Figure 7 show the boxplot of the observed number of SCr measurements, intervention offset, and failure offset, respectively, considering a threshold of 5% graft failure risk in 6 months. The median number of SCr measurements for personalized schedule was lower compared with the fixed schedule (14 vs 29, respectively). Also the inter quartile range (IQR) for number of SCr measurements was lower for personalized schedule compared to as fixed (6 vs 8.5, respectively). The median absolute intervention offset for personalized schedules was 4.5 months, and higher compared with 3.1 months for fixed schedule. Variation of intervention offset was comparable between personalized and fixed schedule (4.1 vs 4.5, respectively). The graft failure offset denoted that in 12% of the times the graft failure was not detected for the test patients when fixed schedule is used, which was comparable with the personalized schedule at a rate of 14%. A more conservative threshold of 2.5% graft failure risk in 6 months resulted in less overshoot of the true graft failure time (6%), however at the expense that it exceeded the true intervention time more often than the fixed schedule (see supplemental file II).

Comment [HPS8]: Considering the length, I've chosen to put this in the supplemental file.

Discussion

The joint model enabled us to study the trajectory of longitudinal SCr and PCR and to specify their association with the risk of graft failure and optimize the screening visits. Nephrologists routinely supervise both the current SCr and PCR level and its increase. We demonstrate that SCr has better discriminative ability for risk of graft failure than PCR. Static baseline clinical data was associated with evolution of SCr, and if included in the model for death-censored graft failure, donor and transplant characteristics were not associated anymore. Nephrologists routinely supervise both the current SCr and PCR level and its increase. Indeed, our results suggest not only the current value of SCr is important but also how rapid the rate of increase to this SCr value was. Also we tested our "one size fits all" screening approach. Our goal was to compare personalized schedule with the currently used fixed schedule, consisting of 20 SCr measurements in the first year after transplantation and hereafter every 3 months. With the joint model, that is inherently patient specific, we show that a personalized screening approach may result in obtaining less SCr measurements while the time to intervene and overcome the risk for graft failure was comparable with the fixed schedule.

Numerous studies have addressed progressively worsening kidney function in a fixed time-window and the risk of graft failure, mostly evaluating linear trajectories of eGFR¹⁹⁻²², some included nonlinearity of progression^{9,23}, but only a few included nonlinear trajectories in a joint model with graft survival which allows dynamic predictions^{24,25}. To understand more precisely how SCr and PCR measurements change over time and how those changes impact graft survival outcomes, joint models efficiently uses all available data. We show which clinical factors associate with evolution of SCr and PCR leading to graft failure, whereas other clinical factors may associate directly with graft failure. We corroborate the findings of Fournier and colleagues who concluded that the graft failure risk depended on both the current value and slope of the SCr.²⁴ We observed that a majority of clinical data at time of transplantation were associated with longitudinal SCr, and less with longitudinal PCR. Once longitudinal SCr and PCR were included in the graft survival model, none of the clinical data at time of transplantation were significantly associated with death-censored graft failure. This could be explained by that the static baseline risk factors are not directly associated with graft failure, which is in line with previous study.²⁴ In our study common risk factors for graft loss such as previous transplantation, cold ischemic time, HLA mismatches, and recipient dialysis vintage were also not significant anymore. This may be partly explained by our study design, that differs in two important aspects. First we were interested in risk factors for the graft, and therefore modelled death-censored graft failure. Secondly, for clinical reasons, we were also interested in predictive ability during the first year, and not hereafter. It may also be that the relatively low number of graft failures increased the risk for type II error in finding predictors for graft failure beside including longitudinal SCr and PCR.

To our knowledge, we firstly compared predictive ability of SCr against PCR, both markers for kidney function decline. We chose to include SCr instead of calculating eGFR to reduce the risk of overcorrection since recipient age, and gender were already included in the longitudinal SCr and PCR

Comment [HPS9]: I'm a bit worried that we didn't include recipient age/gender in our graft failure risk model, to account for creatinine value...

mixed models. Our results suggest that there is no additive predictive value for graft failure of a joint model that includes longitudinal PCR with SCr compared with SCr alone. The KDIGO guidelines provide suggestions for the effective control of serum creatinine and proteinuria, whereas SCr is recommended to measure more frequently than PCR.^{26,27} For measuring urine protein, it is recommended within one month after transplantation as an initial value, and every 3 months during the first year, and annually thereafter.²⁶ For SCr, it is recommended to measure daily within the first week after transplantation, 2-3 times weekly within 2-4 weeks, then once a week within 3 months, then every two weeks within 4-6 months, hereafter monthly within 7-12 months, and every 2-3 months, thereafter.²⁶ Next step is to validate the added predictive value of PCR measurements in other centers, by using dynamic modelling. The variation in SCr measurements in the clinical KDIGO guideline already reflect stability of the graft: if stable, measurements could be reduced by some level. Depending on the clinical transplant data and follow-up measurements, we show a personalized data-driven approach when to plan the next visit.

Comment [HPS10]: What are reasons for this? I can't find literature on this topic...

For the first time to pilot a personalized screenings approach for transplant patients. The framework of joint models allows one to tailor screening to the needs of individual patients and dynamically adapt during follow-up. The fixed and frequent schedules are often burdensome for the transplant patients. Patients who remain relatively stable after transplantation may not require frequent measurement of SCr in the first year. On the other hand, patients for whom the graft decays faster after the first year, a frequent schedule of SCr may be required to check the state of the graft. In this regard, instead of a common fixed schedule for all patients, we propose using a different schedule for every patient. To test whether the personalized schedule is beneficial for the patient, we look at the number of scheduled SCr measurements and at the intervention time at which the 6-month dynamic risk of graft failure of the patient becomes larger than a certain threshold. According to this threshold and the intervention time, we can calculate the time difference between the observed intervention time according to the schedule and the true intervention time (intervention offset), and the time difference between the intervention time and the time of graft failure (graft failure offset). To put it simply, intervention offset calculates 'how much time do we have to prevent graft failure given the schedule?', and graft failure offset calculates 'Are we on time to prevent graft failure given the schedule?'. Given a liberal threshold of 5% risk of graft failure in 6 months, our simulation resulted in a dramatic reduction of number of visits (~50%) while intervention offset and graft failure offset were comparable with the fixed schedule. The next step would be to increase sample size to include more risk factors for graft survival which could facilitate the personalized screening approach. As we included well-known risk factors for graft failure, other biomarkers of interest such as graft histology could be introduced. Finally, our findings have to be externally validated.

Comment [HPS11]: Is this correctly stated in simple language?

Some limitations should be considered as well. Results should be interpreted with care, and externally validated given the relatively low number of events in the survival model. We treated SCr as a surrogate marker for transplant failure, however, other potential causes of graft loss after kidney transplantation exist, including rejection, calcineurin inhibitor toxicity, hypertension, progression of

donor-derived lesions, and recurrence of primary disease. We should note that due to the timing of each next visit depends on the collected longitudinal SCr and PCR of each patient. Therefore we included 13189 SCr measurements and 9616 PCR measurements of only 239 transplant patients in our tertiary referral transplant hospital. Most large transplant registries include monthly or yearly SCr measurements. Also, increasing the sample size may lead to heavy computational burden. As a **quid pro quo**, the joint model, which takes into account the measurement error of longitudinal SCr, overcomes the disadvantages of the Cox model with time-dependent covariates, which has been shown to underestimate the hazard ratios of SCr and graft failure.²⁵

Up to now, only few studies used joint models to evaluate the predictive value of longitudinal measurements of renal function. A greater understanding of renal function trajectory has important implications with regard to the clinical care of transplant patients. SCr has better discriminative ability for risk of graft failure than PCR. Static baseline clinical data was associated with evolution of SCr, and if included in the model for death-censored graft failure, donor and transplant characteristics were not significant associated anymore. Based on the joint model it is possible to select the optimal time point to plan the next measurement, a tailored approach for the patient.

Acknowledgements

...

References

1. Snoeijs MG, Schaubel DE, Hene R, et al. Kidneys from Donors after Cardiac Death Provide Survival Benefit. *Journal of the American Society of Nephrology*. 2010;21(6):1015-1021.
2. Orandi BJ, Luo X, Massie AB, et al. Survival Benefit with Kidney Transplants from HLA-Incompatible Live Donors. *New Engl J Med*. 2016;374(10):940-950.
3. Sorensen VR, Heaf J, Wehberg S, Sorensen SS. Survival Benefit in Renal Transplantation Despite High Comorbidity. *Transplantation*. 2016;100(10):2160-2167.
4. Stegall MD, Gaston RS, Cosio FG, Matas A. Through a Glass Darkly: Seeking Clarity in Preventing Late Kidney Transplant Failure. *Journal of the American Society of Nephrology*. 2015;26(1):20-29.
5. Sellares J, de Freitas DG, Mengel M, et al. Understanding the Causes of Kidney Transplant Failure: The Dominant Role of Antibody-Mediated Rejection and Nonadherence. *Am J Transplant*. 2012;12(2):388-399.
6. Hariharan S, McBride MA, Cherikh WS, Tolleris CB, Bresnahan BA, Johnson CP. Post-transplant renal function in the first year predicts long-term kidney transplant survival. *Kidney international*. 2002;62(1):311-318.
7. Roodnat JL, Mulder PGH, Rischen-Vos J, et al. Proteinuria after renal transplantation affects not only graft survival but also patient survival. *Transplantation*. 2001;72(3):438-444.
8. Streja E, Goldstein L, Soohoo M, Obi Y, Kalantar-Zadeh K, Rhee CM. Modeling longitudinal data and its impact on survival in observational nephrology studies: tools and considerations. *Nephrol Dial Transpl*. 2017;32:77-83.
9. Ferro CJ, Hodson J, Moore J, et al. Bayesian Analysis of Glomerular Filtration Rate Trajectories in Kidney Transplant Recipients A Pilot Study. *Transplantation*. 2015;99(3):533-539.
10. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R.*: CRC Press.; 2012.
11. Rizopoulos D, Taylor JMG, Van Rosmalen J, Steyerberg EW, Takkenberg JJM. Personalized screening intervals for biomarkers using joint models for longitudinal and survival data. *Biostatistics*. 2016;17(1):149-164.
12. Eurotransplant International Foundation. Eurotransplant Manual Chapter 4: ET Kidney Allocation System (ETKAS). 2016;5.5.
13. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: An overview. *Stat Sinica*. 2004;14(3):809-834.
14. Rizopoulos D. The R Package Jmbayes for Fitting Joint Models for Longitudinal and Time-to-Event Data Using MCMC. *Journal of Statistical Software*. 2016;72(7):1-46.
15. Rizopoulos D, Molenberghs G, Lesaffre EMEH. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical J*. 2017;59(6):1261-1276.
16. Rizopoulos D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics*. 2011;67(3):819-829.
17. R Core Team. R: A Language and Environment for Statistical Computing. 2016.

18. Therneau T. A Package for Survival Analysis in S. *version 2.41* 2017; <https://CRAN.R-project.org/package=survival>.
19. Gera M, Slezak JM, Rule AD, Larson TS, Stegall MD, Cosio FG. Assessment of changes in kidney allograft function using creatinine-based estimates of glomerular filtration rate. *Am J Transplant*. 2007;7(4):880-887.
20. Gill JS, Tonelli M, Mix CH, Pereira BJG. The change in allograft function among long-term kidney transplant recipients. *Journal of the American Society of Nephrology*. 2003;14(6):1636-1642.
21. Park WD, Larson TS, Griffin MD, Stegall MD. Identification and Characterization of Kidney Transplants With Good Glomerular Filtration Rate at 1 Year But Subsequent Progressive Loss of Renal Function. *Transplantation*. 2012;94(9):931-939.
22. Kaplan B, Schold J, Meier-Kriesche HU. Poor predictive value of serum creatinine for renal allograft loss. *Am J Transplant*. 2003;3(12):1560-1565.
23. Clayton PA, Lim WH, Wong G, Chadban SJ. Relationship between eGFR Decline and Hard Outcomes after Kidney Transplants. *Journal of the American Society of Nephrology*. 2016;27(11):3440-3446.
24. Fournier MC, Foucher Y, Blanche P, Buron F, Giral M, Dantan E. A joint model for longitudinal and time-to-event data to better assess the specific role of donor and recipient factors on long-term kidney transplantation outcomes. *Eur J Epidemiol*. 2016;31(5):469-479.
25. Asar O, Ritchie J, Kalra PA, Diggle PJ. Joint modelling of repeated measurement and time-to-event data: an introductory tutorial. *Int J Epidemiol*. 2015;44(1):334-344.
26. Kasiske BL, Zeier MG, Chapman JR, et al. KDIGO clinical practice guideline for the care of kidney transplant recipients: a summary. *Kidney international*. 2010;77(4):299-311.
27. Eckardt KU, Kasiske BL. KDIGO Clinical Practice Guideline for the Care of Kidney Transplant Recipients. *Am J Transplant*. 2009;9:S1-S155.

TABLES

Table 1	
<i>Characteristics of the transplant cohort (n=239)</i>	
Donor	Mean (SD) / N (%)
Age (years)	49.7 (12.7)
Gender (Female)	
BMI	25.1 (4.4)
Donor type	
- Living unrelated	38 (15.9%)
- Living related	24 (10.0%)
- DBD	98 (41.0%)
- DCD	79 (33.1%)
Transplant	
Cold ischemia time (hours)	14.8 (8.7)
Panel Reactive Antibodies (%)	
- 0%	181 (75.7%)
- >1%	58 (24.3%)
HLA A, B, DR mismatches	
- 0 mismatch	28 (11.7%)
- 1 mismatch	16 (6.7%)
- 2 mismatches	49 (20.5%)
- 3 mismatches	73 (30.5%)
- 4 mismatches	36 (15.1%)
- 5 mismatches	28 (11.7%)
- 6 mismatches	9 (3.8%)
Recipient	
Age (years)	50.7 (13.1)
Gender (Female)	102 (42.7%)
BMI	25.4 (4.3)
Previous transplantation (yes)	37 (15.5%)
Dialysis vintage (years)	3.7 (3.5)
Diabetes (yes)	38 (15.9%)
Cardiovascular events before Tx (yes)	91 (38.1%)
Number of anti-hypertensives	
- 0	33 (13.8)
- 1	79 (33.1)
- 2	85 (35.6)
≥ 3	42 (17.7)

Table 2: Parameter estimates for the longitudinal model for SCr.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Intercept	5.226	0.080	5.064	5.378	<0.000
Receiver age	-0.063	0.022	-0.107	-0.019	0.010
Donor age	0.083	0.020	0.045	0.119	<0.000
Donor BMI	-0.011	0.021	-0.054	0.028	0.612
Receiver BMI	0.018	0.023	-0.025	0.060	0.420
#HLA mismatches between donor and recipient	0.020	0.022	-0.022	0.065	0.342
Panel reactive antibody percentage	0.048	0.027	-0.008	0.100	0.082
#Anti-hypertensive medicaments	0.040	0.020	0.001	0.080	0.048
Cold ischemia time	0.029	0.035	-0.039	0.102	0.390
#Days on dialysis before transplant	0.015	0.029	-0.042	0.071	0.580
Receiver gender: Male	0.197	0.042	0.111	0.276	<0.000
Previous transplant: Yes	0.016	0.064	-0.115	0.141	0.786
Donor gender: Male	0.053	0.042	-0.027	0.136	0.198
Delayed graft function: Yes	0.118	0.049	0.025	0.216	0.006
Diabetes Mellitus: Yes	-0.103	0.059	-0.217	0.012	0.076
Cardiovascular events before transplantation: Yes	-0.047	0.043	-0.129	0.044	0.272
Deceased donor: Yes	0.163	0.082	0.004	0.313	0.044
Spline: visit time [0.039, 0.082] years	-0.440	0.041	-0.517	-0.358	<0.000
Spline: visit time [0.082, 0.219] years	-0.182	0.053	-0.284	-0.081	<0.000
Spline: visit time [0.219, 1] years	-0.545	0.081	-0.712	-0.395	<0.000
Spline: visit time [1, 6] years	0.007	0.083	-0.155	0.176	0.946
σ	0.190	0.001	0.187	0.192	

Table 3: Parameter estimates for the longitudinal model for PCR.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Intercept	3.731	0.179	3.398	4.083	<0.000
Receiver age	0.030	0.052	-0.066	0.138	0.604
Donor age	0.209	0.047	0.118	0.301	<0.000
Donor BMI	-0.019	0.051	-0.121	0.084	0.716
Receiver BMI	-0.116	0.050	-0.219	-0.021	0.014
#HLA mismatches between donor and recipient	-0.013	0.049	-0.112	0.086	0.776
Panel reactive antibody percentage	0.047	0.061	-0.066	0.166	0.446
#Anti-hypertensive medicaments	0.056	0.047	-0.03	0.147	0.208
Cold ischemia time	0.062	0.082	-0.097	0.211	0.468
#Days on dialysis before transplant	0.006	0.066	-0.120	0.134	0.952
Receiver gender: Male	-0.026	0.094	-0.207	0.166	0.798
Previous transplant: Yes	0.035	0.149	-0.241	0.332	0.816
Donor gender: Male	0.114	0.096	-0.079	0.303	0.228
Delayed graft function: Yes	0.043	0.118	-0.174	0.275	0.740
Diabetes Mellitus: Yes	0.153	0.135	-0.124	0.396	0.256
Cardiovascular events before transplantation: Yes	-0.016	0.106	-0.221	0.199	0.890
Deceased donor: Yes	0.144	0.193	-0.246	0.509	0.462
Spline: visit time [0.039, 0.082] years	-0.821	0.090	-0.989	-0.638	<0.000
Spline: visit time [0.082, 0.219] years	-0.578	0.131	-0.838	-0.304	<0.000
Spline: visit time [0.219, 1] years	-0.898	0.160	-1.218	-0.587	<0.000
Spline: visit time [1, 6] years	0.460	0.234	0.015	0.927	0.036
σ	0.479	0.004	0.472	0.486	

Table 4: Relative risk sub-model estimates for mean and 95% credible interval.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Previous transplant: Yes	0.305	0.339	-0.099	0.986	0.352
#HLA mismatches between donor and recipient	0.048	0.093	-0.114	0.269	0.620
Cold ischemia time	-0.051	0.105	-0.277	0.133	0.644
#Days on dialysis before transplant	-0.013	0.102	-0.251	0.178	0.934
log PCR	0.145	0.125	-0.056	0.431	0.188
Slope(log PCR)	0.021	0.058	-0.076	0.145	0.828
log SCr	1.599	0.241	1.067	2.063	<0.000
Slope(log SCr)	0.203	0.123	-0.017	0.443	0.082

Table 5: Area under curve characteristics for the JMs fitted to the kidney transplant data set.

Biomarkers	Year 0.5	Year 1	Year 1.5	Year 2	Year 2.5	Year 3
Both SCr and PCR	0.845	0.927	0.915	0.828	0.953	0.686
Only SCr	0.801	0.901	0.918	0.866	0.957	0.692
Only PCR	0.844	0.858	0.755	0.640	0.825	0.473

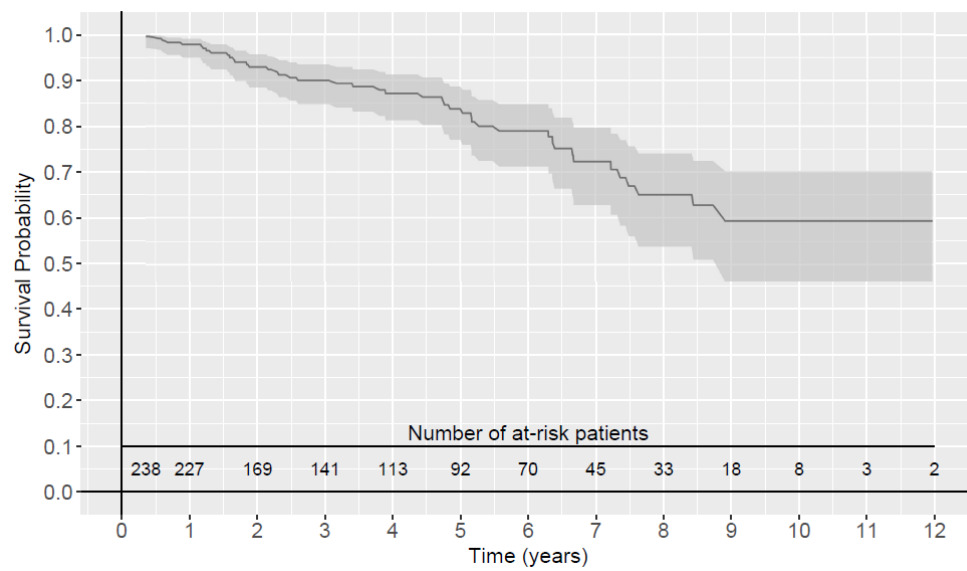


Figure 1: Death-censored graft survival and 95% CI using Kaplan Meier.

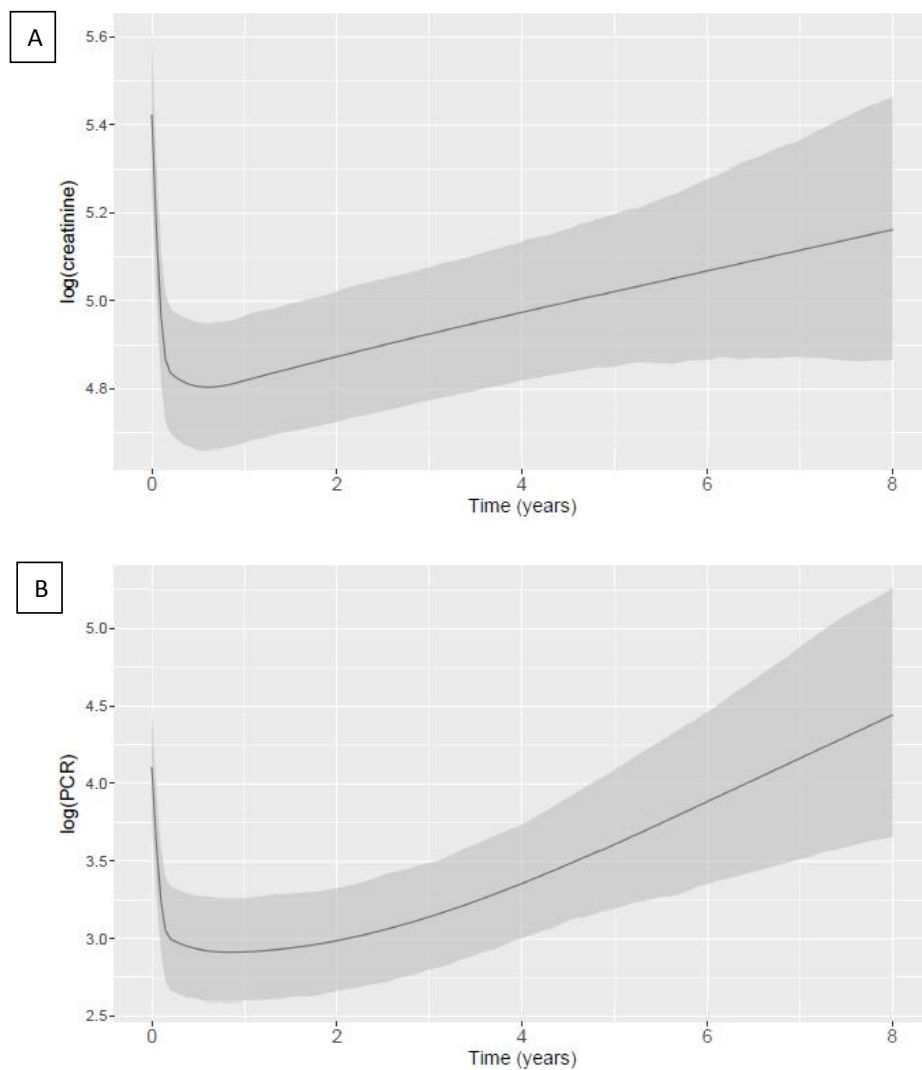


Figure 2: Fitted longitudinal evolution of SCr (A) and PCR (B) with 95% CI, corresponding to a female recipient aged 50.7 years, BMI 25.4, first transplantation, no diabetes, no history of any cardiovascular events, of a living female donor aged 49.7 years, with 3 mismatches on HLA A, B, DR, with 14.8 hours of CIT, 4.8% of Panel Reactive Antibodies, and 3.7 years of dialysis vintage.

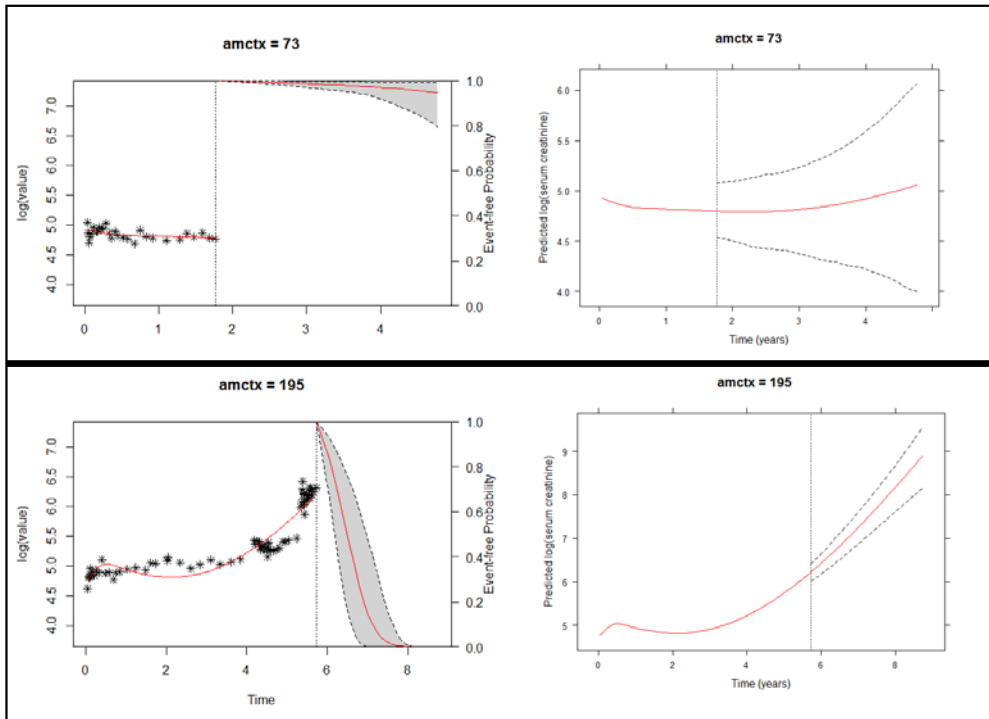


Figure 3: Dynamic predictions of death-censored survival probabilities and dynamic predictions of log(creatinine) up to 3 years after loss of follow up.

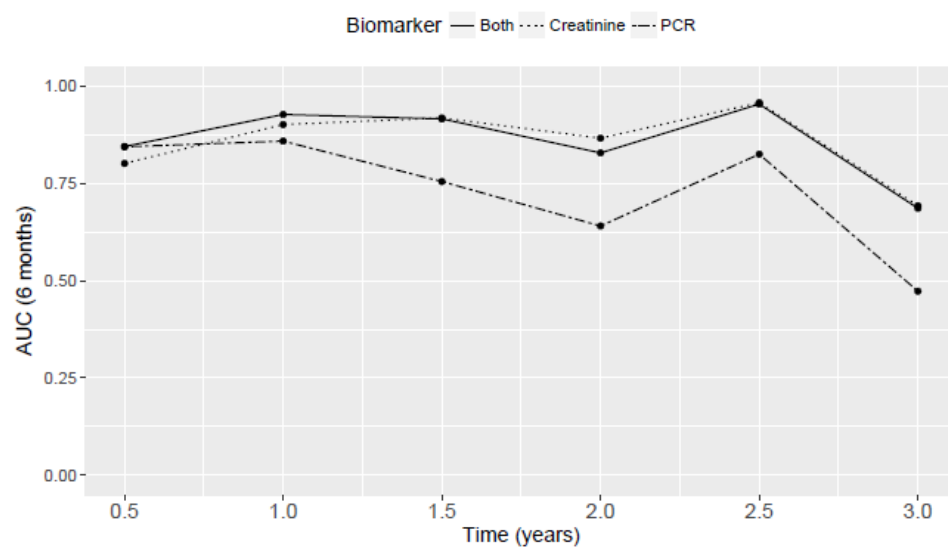


Figure 4: Time-dependent area under the receiver operating curves (AUC) for the joint models.

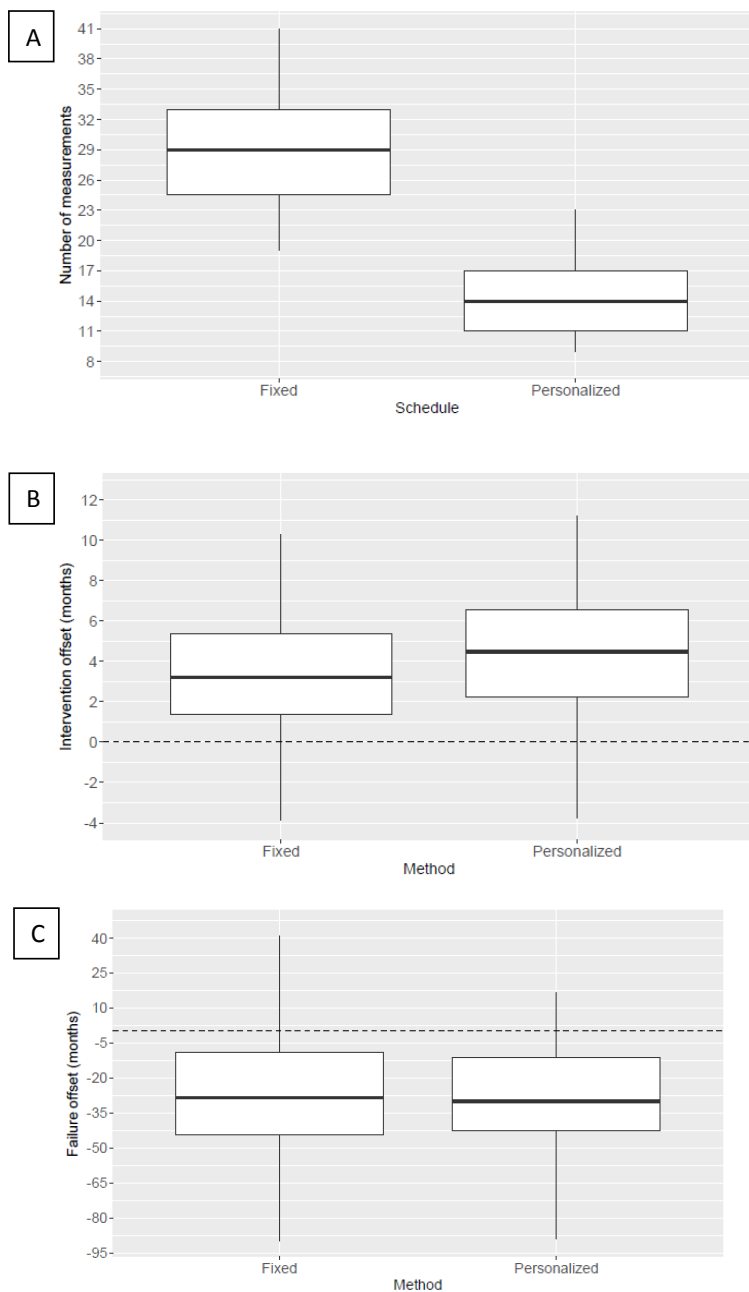


Figure 5: Box plots of the number of scheduled SCr measurements (A), intervention offset (B), and the graft failure offset (C). Fixed schedules were compared with personalized schedules, depended on the dynamic predictions of the joint model. The threshold was set at 5% risk of graft failure per 6 months. The zero offset mark (for B and C) is displayed with the dashed line.

