Department of Biostatistics
Erasmus University Medical Center

Internal data analysis Report

Title	PRIAS
Version	1
Date	10 February 2017
Author	
Collaborators	

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 a preliminary analysis of the prostate cancer data set from the PRIAS study. The original data set as of December 7 2016 contains information about 6039 patients who were inducted in the study under the conditions of active surveillance(AS). To detect the progression of prostate cancer, PSA, DRE and Gleason scores were measured on these patients based on a fixed schedule. The age of the patients at baseline was also recorded. Although Gleason scores based on biopsy are the gold standard for detection of prostate cancer progression, they are painful and have side effects lasting months. The purpose of the we present in this report is to know the relationship between PSA, DRE and Gleason scores and see check if the former 2 can be used as substitutes for Gleason.

1.1 Gleason score as indicator for progression

There are two approaches to find the solution for the aforementioned goal. The first way is to do a multivariate longitudinal analysis to find the strength of the association between the outcomes. The second one, which we present in this report, is to consider the Gleason score as indicator for a time to event outcome and check the association between PSA/DRE and the time the event outcome via a joint model. The event of interest here is Gleason > 6, which may be seen as progression of cancer.

It is important to note that Gleason scores are prone to a downward measurement error. i.e. Gleason scores may be measured lesser than what they really are. The vice versa doesn't happen though. Secondly, a Gleason score 7 can be obtained in two ways: 7=3+4 (3 for most common cell type and 4 for second most common cell type) and 7=4+3. The former is not considered as risky as the latter. However from the data set it is not possible to know if the Gleason score 7 was formed as 3+4 or 4+3. Lastly, since Gleason scores are measured on a fixed schedule, the corresponding time to event outcome is prone to interval censoring. However in this preliminary analysis we assume that Gleason scores are devoid of the aforementioned problems.

1.2 Selection of patients

In the analysis presented here, we considered only 5131 patients out of the 6039 available. The rest of the patients were not considered because of the following reasons.

- Patients having no other information other than the ID.
- Patients having age less than 5.
- Patients not having any information on follow ups.
- Patients who discontinued but do not have any date of discontinuation.
- Patients not having reasons of discontinuation.
- Patients with incomprehensible reasons of discontinuation.
- Patients with Gleason score > 6 at the first visit.

Further we also removed all those observations where:

- PSA/DRE/Gleason scores were available but the date of measurement was not.
- Date of measurement is available but none of the 3 scores are available.
- Measurements were dummy.
- Gleason scores were 0 or 1.
- PSA/DRE were measured on dates after the date of first occurrence of Gleason > 6.

Since the event of interest is Gleason > 6, the patients who never had a Gleason score greater than 6, were considered to be censored. The last follow up date for these patients was considered to be the date of censoring. It is important to note that the original study protocol does not consider Gleason > 6 as progression and in many cases patients with a Gleason score 6 are given treatment

and patients with Gleason > 6 are on AS. This is further discussed in section 2 and section 5.3. We have dichotomized DRE scores (T1c and >T1c), as some of the categories of DRE scores had too less patients and besides our software doesn't support categorical outcomes with more than 2 levels at this moment.

1.3 Various transformations of available data

In this analysis we did transformations of available data in multiple scenarios. For e.g. in the longitudinal analysis of PSA measurements we used log(PSA + 1) measurements instead of the raw data. The log transformation was done after considering that certain patients had very large PSA scores and thus suggested that the underlying distribution is right skewed. Secondly a lot of patients had PSA score 0, and thus we added a constant 1 to all PSA scores before log transformation. We found the same transformation in literature as well [McGreevy et al., 2006, Sène et al., 2016].

For the variable Age in the Cox model, we wanted to also consider the quadratic effect. However doing so led many numerical instabilities while performing the joint model analysis. To avoid these pitfalls we had to center the Age variable around age 70 years.

Lastly, we used piecewise regression techniques for modeling the evolution of log (odds) for DRE and we used splines for log(PSA+1) score. We tried models with linear evolutions over time as well, however they lead to a very sharp evolution of the responses, which did not match the observed evolutions. Further, in case of PSA we tried various selection of knots and found that more or less the same evolution was observed. Although the coefficients for splines are hard to interpret, we have presented graphs to fitted evolutions to make them interpretable.

2 Descriptive Statistics

Out of the 5131 subjects who were considered for analysis, only 670 had Gleason > 6 at least once. Table 1 shows the number of subjects in the study by their treatment/event categories. As we can see not everyone who was given a treatment based on protocol had an increase in DRE or had a Gleason score reclassification. We can also see that many of the subjects who had Gleason reclassification or had an increase in DRE are still on AS. For the latter scenario, we lack the clinical knowledge to give a reason for continuation on AS. Nevertheless for the purpose of the current analysis, the definition of event used is Gleason > 6 and we ignore the event information used in the study protocol.

	Died	Treatment	WW	LFU	Anxiety, On-Req., Other	OnAS	Total
Gleason > 6 at least once	2	482	27	1	42	116	670
Increase in DRE at least once	15	277	28	16	63	477	876
Total	63	1126	179	74	372	3317	5131

Table 1: Number of subjects by event category.
WW: Watchful waiting, LFU: Lost to follow up, On-Req: On request.

3 Time to event analysis of disease progression

To understand how to model the time to event outcome, we began with a non parametric analysis for time to Gleason > 6. Figure 1 shows the corresponding Kaplan Meier survival curves for 3 age groups, Age ≤ 67 years, Age > 73 years and Age between 67 and 73 years. We can see that there is a difference between the survival curves of the 3 age groups after the first year. Particularly the lowest age groups seems to have higher survival probability compared to other age groups at all times. To confirm the effect of Age, we fitted a Cox model for time to progression, with both Age and its quadratic effect as covariates. We found both covariates to be significant at 95% level of significance, however the effect sizes were so small that for all practical purposes the overall effect of Age was ignorable up to a Age difference of 25 years.

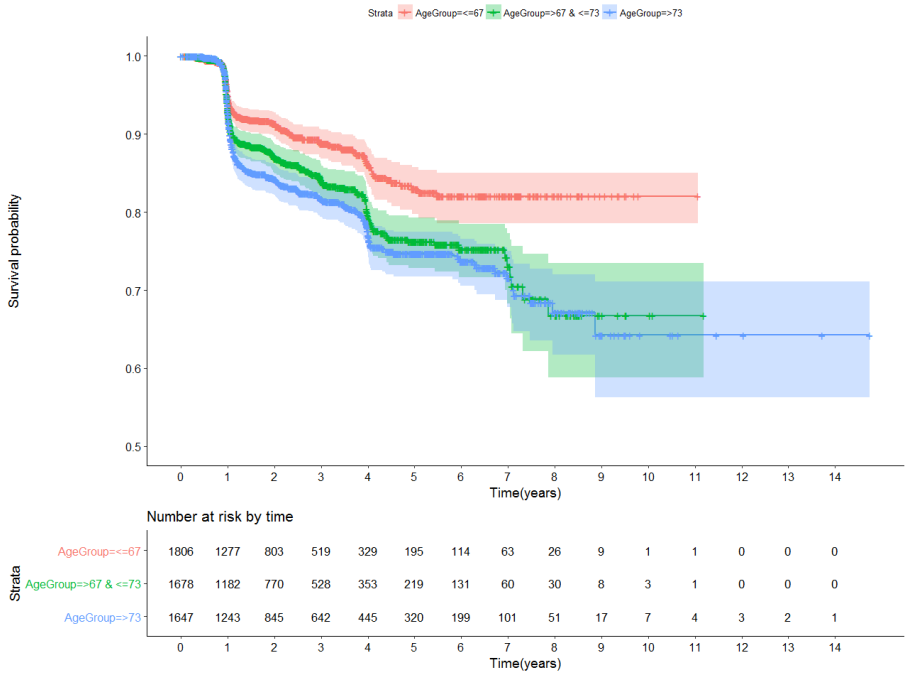
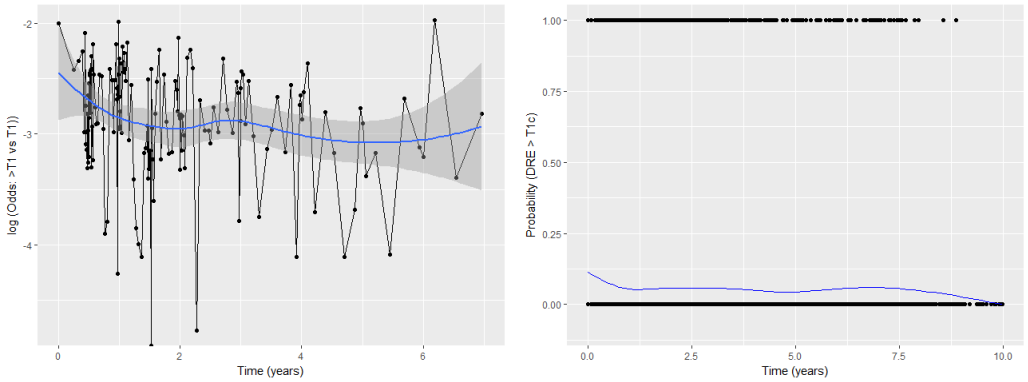


Figure 1: Survival probability for 3 age groups. Age $\leq$ 67 years, Age $>$ 73 years and Age between 67 and 73 years.

4 Longitudinal analysis of PSA and DRE

For the univariate longitudinal analysis of DRE we first plotted the observed evolutions of the log odds based on a categorized time scale. This was done because we modeled the log odds for both of these responses using Generalized linear mixed models (GLMM). The observed evolutions for log(odds) are shown in Figure 2a. Correspondingly, the observed evolution of Probability(DRE $>$ T1c) is shown in Figure 2b. Based on these plots we decided to use segmented evolutions with 2 segments in total to model the evolution for DRE outcome. The choice of knot between the segments was based on opinion of experts and Figure 2. The knot for evolution of log odds of DRE score was selected at 1 year. For the random effects we only considered the random intercept in this preliminary analysis.



(a) Observed evolution of log(odds: $>$T1c against T1c).

(b) Observed evolution of Probability(DRE $>$ T1c).

Figure 2: Observed evolutions of log(odds: $>$T1c against T1c) and Prob(DRE $>$ T1c) for DRE scores. The time scale in the log(odds) plot is categorized. Thus log(odds) at time point t are calculated by taking all observations between time t and $t-1$.

To understand the evolution of log(PSA+1) scores, we plotted the observed evolutions of

log(PSA+1) scores after adjusting for Age (Figure 3). However after checking various subsamples of patients (individual as well as marginal over the entire subsample) we could not see a clear trend visually. So instead we then fitted various models to come up with a good fit. After comparing various models and considering the computational restrictions we chose a longitudinal model with a spline with 3 internal knots to model fixed effect and with 1 internal knot to model random effect of time.

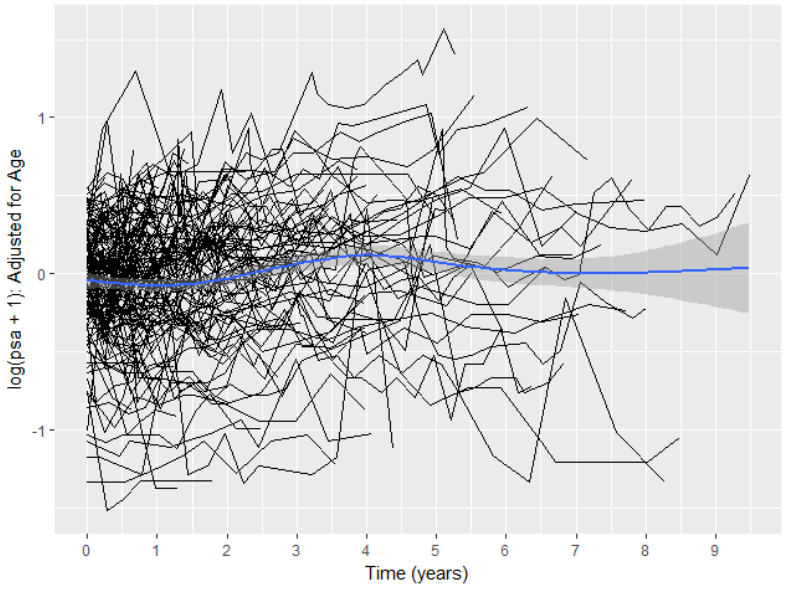


Figure 3: Observed evolutions for log(PSA + 1) measurements after adjusting for Age of the patients. A random sample of 150 patients is used.

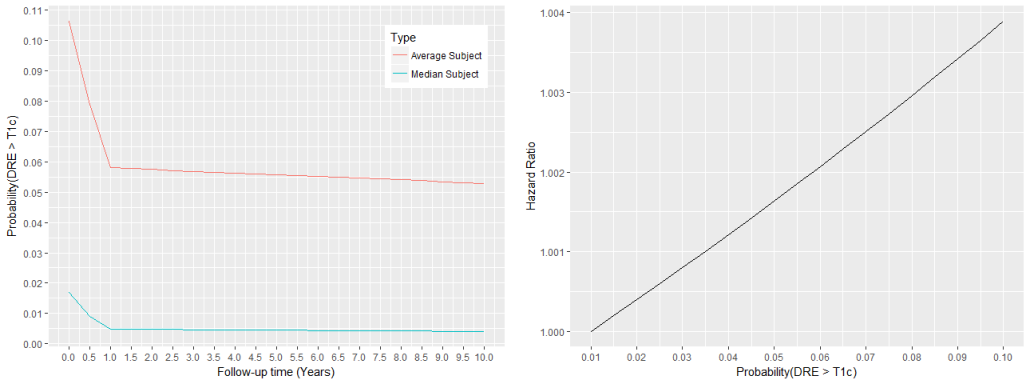
5 Joint model for analysis of PSA, DRE and Gleason

In this preliminary analysis, we present 3 separate joint models. The first two are with univariate longitudinal models for PSA and DRE and the last one is having a multivariate longitudinal model for PSA and DRE. The univariate models were fitted with mvJointModelBayes API instead of jointModelBayes because the jointModelBayes API did not lead to models with convergence for MCMC outcomes. Convergence was evaluated graphically using traceplots.

5.1 DRE score

In the joint model for DRE score we used the Odds of DRE score $>$ T1c against DRE score T1 as the association parameter. For the resulting model the MCMC outcomes did not show any signs of non convergence for any of the parameter estimates. The resulting parameter estimates are shown in Table 2 and Table 3. The fitted evolution of Probability(DRE $>$ T1c) is shown in Figure 4a. The average evolution a patient of Age 70 years is shown in red. The evolution for a median subject of Age 70 is shown in blue. This subject is the one with with 98.3% chance of having a DRE = T1c at the time of induction in the study. These fitted evolutions are comparable to the observed evolution in Figure 2b. The effect of Age is also significant, and for a particular patient, if Age at the time of induction was 8.7 years more than what they had, then their odds of having a DRE score $>$ T1c would've been twice. The odds would've been 4 times if the Age of induction was 17.4 years more than what they had. We further verified that this particular odds ratio increases more or less linearly with increase in Age.

The association of Odds of DRE score with survival outcome (Gleason $>$ 6) is not significant. This is also evident in the hazard ratio plot shown in Figure 4b. This plot shows the hazard ratio for the same patient, had it been having a different probability of DRE score $>$ T1c. As we can see, the hazard of having a progression increases trivially with change in probability. The combined



(a) Fitted evolution of probability of having a DRE score > T1c over a period of 12 months, for a patient of Age 70 years.

(b) Hazard ratio for various Prob(DRE > T1c) against Prob (DRE > T1c) = 0.01. Only valid for the same subject at the same time point.

Figure 4: Fitted evolution of probability of having a DRE score > T1c over time and Hazard ratio for various Prob(DRE > T1c) against Prob (DRE > T1c) = 0.01.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Intercept	-9.651	0.708	-11.084	-8.290	<0.000
Age	0.080	0.009	0.062	0.098	<0.000
visitTimeYears	-1.255	0.093	-1.437	-1.074	<0.000
(visitTimeYears - 1) × I(visitTimeYears >1)	1.235	0.109	1.029	1.448	<0.000

Table 2: Longitudinal submodel estimates for joint model with DRE score outcome. Outcome modeled is log odds of DRE score >T1c against T1c.

effect of Age and its quadratic effect although significant, is such that for all practical purposes Age can be ignored. It is important to note that these comparisons are made for the same patient and at the same time.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Age - 70	0.035	0.008	0.021	0.049	<0.000
(Age - 70) × (Age - 70)	-0.002	0.001	-0.003	-1.313 $\times 10^{-4}$	0.038
Odds (>T1c against T1c)	0.038	0.026	-0.012	0.086	0.120

Table 3: Survival submodel estimates for joint model with DRE score outcome.

5.2 PSA score

In the joint model for PSA score, the MCMC outcomes did not show any sign of non convergence for any of the parameter estimates including the components of variance covariance matrix. The resulting parameter estimates are shown in Table 4 and Table 5. Since the longitudinal evolution of log(PSA + 1) is fitted with non-linear terms, the interpretation of the coefficients corresponding to time is not straightforward. However in Figure 5 we present the fitted evolution of PSA over a period of 10 years for a patient who is 70 years old. It can be seen that the PSA scores go down in the first 1 year and then steadily increase over the next 9 years. Since the model has only additive terms, this evolution remains same for all patients. The effect of Age only affects the baseline PSA score. However it is so small that it can be ignored for all practical purposes.

For the survival submodel it can be seen that only the rate of change of log(PSA + 1) is associated with time to progression (Gleason > 6). The effect is quite strong and if at any given time point the rate of change of log(PSA + 1) becomes 2 times then the hazard of progression (Gleason > 6) becomes almost 98 times more. This is valid under the condition that instantaneous

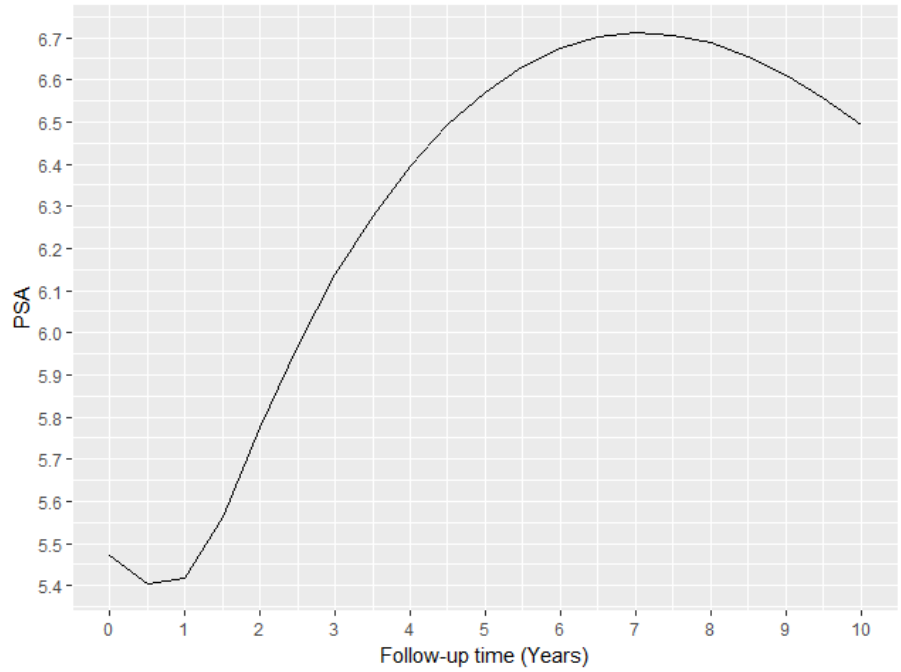


Figure 5: Fitted evolution of PSA over a period of 10 years, for a patient who was inducted in AS at the Age of 70 years.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Intercept	1.867	0.006	1.855	1.879	<0.000
Age - 70	0.002	0.001	0.001	0.004	<0.000
(Age - 70) × (Age - 70)	-3.455 $\times 10^{-4}$	7.199 $\times 10^{-5}$	-4.796 $\times 10^{-4}$	-2.015 $\times 10^{-4}$	<0.000
Spline: visitTimeYears[0, 1]	0.073	0.007	0.059	0.086	<0.000
Spline: visitTimeYears[1, 2]	0.273	0.020	0.234	0.314	<0.000
Spline: visitTimeYears[2, 4]	0.110	0.073	-0.026	0.251	0.134
Spline: visitTimeYears[4, 15]	0.033	0.139	-0.225	0.299	0.802

Table 4: Longitudinal submodel estimates for joint model with log(PSA + 1) score outcome.

value of PSA scores are same for the comparison. Lastly, for the effect of Age on hazard we can say that it can be safely ignored for all practical purposes.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Age - 70	0.037	0.008	0.021	0.052	<0.000
(Age - 70) × (Age - 70)	-0.001	0.001	-0.003	1.960 $\times 10^{-4}$	0.090
log(PSA + 1)	0.164	0.140	-0.119	0.430	0.232
Slope: log(PSA + 1)	4.584	0.717	3.285	6.070	<0.000

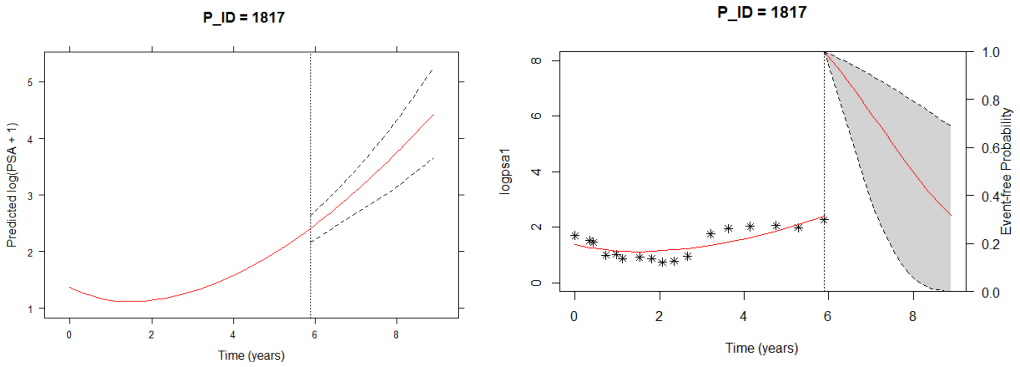
Table 5: Survival submodel estimates for joint model with log(PSA + 1) score outcome.

5.3 Dynamic Predictions

Using the information from the fitted joint model, we computed the Dynamic predictions for the longitudinal outcome log(PSA + 1) and survival probabilities of various patients. It is important to note that the dynamic predictions of survival probability do not use the dynamic predictions of longitudinal outcome, and are entirely based on observed longitudinal data for the corresponding patient.

5.3.1 Patient ID: 1817

Patient with ID 1817, never turned up for any biopsy after the first one at the time of induction in the study. However the patient turned out regularly for all PSA measurements. The dynamic predictions of longitudinal and survival for this patient are shown in Figure 6. We can see that the observed PSA scores increase quite rapidly as the latter follow up times. Correspondingly based on our model the survival probability declines rapidly in the 3 years after the last follow up time. This indicates that there is a high chance that the Gleason scores will become > 6 in the next 3 years. Furthermore the longitudinal outcome $\log(\text{PSA} + 1)$ is also predicted to increase rapidly.



(a) Dynamic predictions for the longitudinal outcome.

(b) Dynamic predictions of the survival probability.

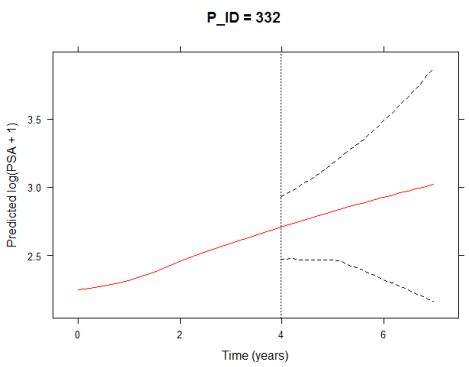
Figure 6: Dynamic predictions for the longitudinal and time to event outcome of Patient 1817.

5.3.2 Patient ID: 332

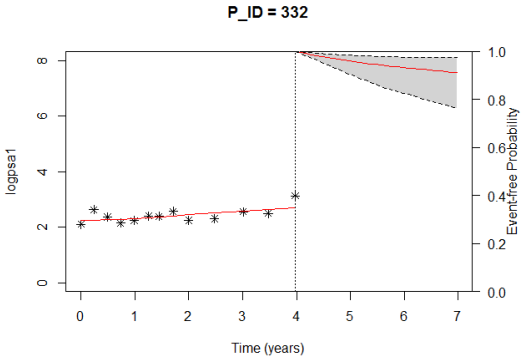
Patient with ID 332, turned up 4 times in total for biopsy but never had a Gleason score > 6. The patient never had a DRE more than T1c. However the PSA for the patient almost tripled in the 4 years he was on AS. This patient was given a treatment for cancer on the basis of study protocol (we ignore this since the event of interest if Gleason > 6). However after checking the study protocol of PRIAS we could not justify the reason of treatment given as "based on study protocol". The dynamic predictions of longitudinal and survival for this patient are shown in Figure 7. We can see that the observed PSA scores increase with follow up times, however the rate of increase is more or less constant. Correspondingly based on our model the survival probability declines slowly in the 3 years after the last follow up time. This indicates that there is a small chance that the Gleason scores will become > 6 in the next 3 years. Although now that the treatment has been given, this result may not be of interest.

5.3.3 Patient ID: 4324

Patient with ID 4324 had a more or less stable PSA score over time. The patient had a DRE score T2a (a,b) for the first 3 DRE visits however for the latter 3 it decreased to T1c. Furthermore the Gleason score was measured only once at the induction in AS. As expected from these optimistic observations, this patient was not given a treatment. The dynamic predictions of longitudinal and survival for this patient are shown in Figure 8. We can see that the PSA score is predicted to decrease for this patient in the 3 years since the last follow up time. Also the dynamic predictions for time to event outcome show that there is little to no chance that the patient will have a Gleason score > 6 in the 3 years since last follow up.

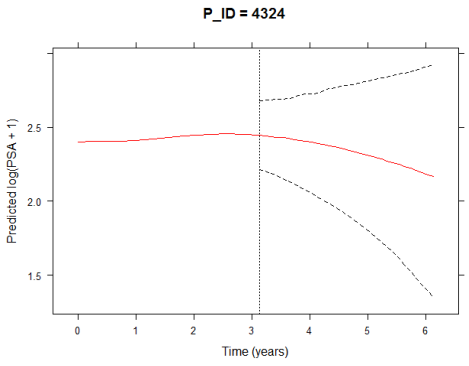


(a) Dynamic predictions for the longitudinal outcome.

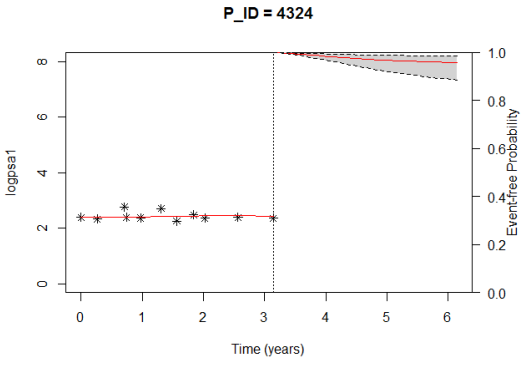


(b) Dynamic predictions of the survival probability.

Figure 7: Dynamic predictions for the longitudinal and time to event outcome of Patient 332.



(a) Dynamic predictions for the longitudinal outcome.



(b) Dynamic predictions of the survival probability.

Figure 8: Dynamic predictions for the longitudinal and time to event outcome of Patient 4324.

References

[McGreevy et al., 2006] McGreevy, K., Rodgers, K., Lipsitz, S., Bissada, N., and Hoel, D. (2006). Impact of race and baseline psa on longitudinal psa. *International journal of cancer*, 118(7):1773–1776.

[Sène et al., 2016] Sène, M., Taylor, J. M., Dignam, J. J., Jacqmin-Gadda, H., and Proust-Lima, C. (2016). Individualized dynamic prediction of prostate cancer recurrence with and without the initiation of a second treatment: Development and validation. *Statistical methods in medical research*, 25(6):2972–2991.