Harvard University

Harvard University Biostatistics Working Paper Series

Year 2017 *Paper* 214

Studying the Optimal Scheduling for Controlling Prostate Cancer under Intermittent Androgen Suppression

Sunil K. Dhar* Hans R. Chaudhry† Bruce G. Bukiet‡

Zhiming Ji** Nan Gao†† Thomas W. Findley‡‡

*Center for Applied Mathematics and Statistics and the Department of Mathematical Sciences, New Jersey Institute of Technology, dhar@njit.edu

†Vetha Center of Transdisciplinary Studies, NJII, NJIT, hans.chaudhry@consciousnessscience.org

‡Center for Applied Mathematics and Statistics and the Department of Mathematical Sciences, New Jersey Institute of Technology, bruce.g.bukiet@njit.edu

**Department of Mechanical and Industrial Engineering, NJIT, zhiming.ji@njit.edu

††Department of Biological Sciences, Rutgers, The State University of New Jersey, ngao@newark.rutgers.edu

‡‡Cancer Institute of New Jersey, Physical Medicine, New Jersey Medical School, Rutgers, the State University of New Jersey, tom.findley@gmail.com

Studying the Optimal Scheduling for Controlling Prostate Cancer under Intermittent Androgen Suppression

Sunil K. Dhar, Hans R. Chaudhry, Bruce G. Bukiet, Zhiming Ji, Nan Gao, and Thomas W. Findley

Abstract

This retrospective study shows that the majority of patients' correlations between PSA and Testosterone during the on-treatment period is at least 0.90. Model-based duration calculations to control PSA levels during off-treatment are provided. There are two pairs of models. In one pair, the Generalized Linear Model and Mixed Model are both used to analyze the variability of PSA at the individual patient level by using the variable "Patient ID" as a repeated measure. In the second pair, Patient ID is not used as a repeated measure but additional baseline variables are included to analyze the variability of PSA.

Studying the Optimal Scheduling for Controlling Prostate Cancer under Intermittent Androgen Suppression

Sunil K. Dhar, PhD[1]; Hans R. Chaudhry, PhD[2]; Bruce G. Bukiet, PhD[1]; Zhiming Ji, PhD[3]; Nan Guo, PhD[4]; and Thomas W. Findley, MD, PhD[5]

Abstract This retrospective study shows that the majority of patients' correlations between PSA and Testosterone during the on-treatment period is at least 0.90. Model-based duration calculations to control PSA levels during off-treatment are provided. There are two pairs of models. In one pair, the Generalized Linear Model and Mixed Model are both used to analyze the variability of PSA at the individual patient level by using the variable "Patient ID" as a repeated measure. In the second pair, Patient ID is not used as a repeated measure but additional baseline variables are included to analyze the variability of PSA.

[1]Center for Applied Mathematics and Statistics and the Department of Mathematical Sciences, New Jersey Institute of Technology (NJIT)

[2]Vetha Center of Transdisciplinary Studies, NJII, NJIT

[3]Department of Mechanical and Industrial Engineering, NJIT

[4]Department of Biological Sciences, Rutgers, The State University of New Jersey (RTSUNJ)

[5]Cancer Institute of New Jersey, Physical Medicine, New Jersey Medical School, Rutgers, the State University of New Jersey

Corresponding author: Sunil K. Dhar, Department of Mathematical Sciences, 323 Martin Luther King Junior Boulevard, Newark, New Jersey 07102; dhar@njit.edu

Running title: Optimal Scheduling PCa Treatment

Keywords: PSA; Testosterone; Predicted Duration; Generalized Linear Model; Mixed Model.

Introduction According to the American Cancer Society's estimates for prostate cancer in the United States in 2015, there are about 220,800 new cases of prostate cancer and about 27,540 deaths from prostate cancer each year. It is the second most fatal type of cancer for males, behind only lung cancer. One out of seven men will get prostate cancer during his lifetime and one in 38 will die of this disease. Available from URL: (http://www.cancer.org/cancer/prostatecancer/detailedguide/prostate-cancer-key-statistics)

[accessed May 4, 2015]

Prostate Cancer (PCa) represents a third of all cancer incidence in men – six-fold that of leukemia and lymphoma. Most cases are indolent, referring to the number of elderly men who die with cancer rather than die from it. Huggins, a urological surgeon, reasoned "that rather than eliminate the male hormones, what if one tricked the cancer into thinking that the body was 'female' by suppressing the effect of testosterone?" This led to synthesizing analogues of estrogen, the female hormone, which resulted in the most widely used drug, DES, or Premarin, in order to stop the production of testosterone in patients with prostate cancer, Mukherjee (2010). "PCa cells at an early stage also have a dependence on androgen, androgen ablation results in the degeneration of prostate tumor. The major treatment for advanced PCa has been endocrine (hormone) therapy. If androgen deprivation (AD) is overly prolonged (continuous androgen suppression, CAS), the AD cells change to androgen – independent (AI) cancer cells which are resistant to hormone therapy. It is possible to suspend androgen suppression by simply stopping the administration of the hormone based drugs. Building on this concept, intermittent androgen suppression (IAS) was proposed, Hirata, et al (2012). In IAS treatment, when an upper threshold value of PSA is reached, therapy is started while when the PSA level decreases to a lower threshold level, treatment is stopped. PCa can be treated by one or more of the following methods, depending upon the stage of PCa: 1.Surgery, 2. Radiation Therapy, 3. Hormone Therapy, 4. Chemotherapy.

It is noted that hormone therapy cannot cure PCa. Instead, it slows the growth of PCa. It is mostly used in last stages, i.e., C&D. It decreases the production of testosterone. Since we are investigating IAS treatment in this paper, it is worth knowing the effectiveness of this treatment. The efficacy of IAS was demonstrated in animals, Bruchovsky et al. (1990) and Akakura et al.

(1993). The clinical efficacy of IAS for human patients has been studied by many research groups, Abrahamsson (2010). One hundred and three patients after radiation therapy for locally advanced PCa, eligible for IAS were treated for 6 years. The time off treatment averaged 35% of the total time cycle during an average of 5 cycles. Prostate volume was reduced by 40% in cycle 1 and by 34% in cycle 2, but there was no decrease in subsequent cycles Bruchovsky et al. (2006). It was reported by Hussain M, et al. (2013) that a total of 1535 patients were analyzed to compare IAS versus CAS therapy in PCa. Of these, 765 were randomly assigned to CAS and 770 to IAS. The mean follow up period was 9.8 years. The IAS group was associated with better erectile function and mental health at 3 months but not thereafter. There was a 20% greater risk of death with IAS therapy. However, no significant inferiority of IAS therapy was observed. IAS therapy resulted in small improvements in quality of life. These conclusions should be taken into consideration in recommending IAS or CAS therapy to patients. It is also observed that IAS treatment on PCa shows remarkable results at the beginning of treatment, but cancer cells frequently grow during long term CAS therapy. IAS is effective for some patients but not others. To understand these problems, mathematical models have been developed to explain the mechanism of IAS and CAS. A review of various mathematical models is made, Hirata, et al (2012). Out of all these models, the piecewise linear model is preferred. It compares the prediction of the outcome of IAS and CAS hormone therapy and predicts the actual quantitative changes in prostate specific antigen (PSA), which is a tumor marker, Hirata et al. (2010-1). However, no optimal scheduling for IAS was done in this paper. The method of optimal scheduling was explained later by Hirata et al. (2010-2). This optimal scheduling determines a recommended duration for the 'on-treatment period' and the 'off-treatment period' such that PSA declines.

Many statistical models in prostate research can be found in the literature. Some recent ones used are tests for trend and multivariable logistic regression analyses, Bhindi et al. (2015). The usual two-sided statistical tests and area under the curve methods are used by Bryant et al. (2015). Multivariate random-effects meta-analysis models are used by Chen et al. (2015). Survival analysis based methods in particular Kaplan-Meier analysis and one variable and multivariable Cox proportional hazards models are also used by Moschini et al. (2016). A statistic shape model was used recently, Tao et al. (2015). Although many statistical models exist in the literature, to the best of the knowledge of the present authors, no statistical model investigating the optimal scheduling of on- and off-treatment duration has yet been developed. Mathematics based models to study the problem of optimal scheduling has been looked at by Hirata et al. (2010-2) but this manuscript is the first of its kind using the specified statistical models to study this issue. For illustration, consider a prostate cancer patient who has been under the treatment of an oncologist. The oncologist puts the patient under treatment, say, for about two months. The patient reports to the oncologist after two months and his observed PSA has, say, come down below the normal level. Thus, the oncologist decides to discontinue his treatment for say two or three months due to improvement of the patient's quality of life. However, it is quite likely that during this off treatment period his PSA rises significantly due to lack of continuous monitoring and therapy. Therefore, in this manuscript, we want to find the optimal duration for on- and off-treatment periods so that PSA value does not go much beyond the final PSA reading during on-treatment.

Methods. We use the Generalized Linear Modeling and Mixed Modeling approaches to build our statistical models. The Generalized Linear Model is the generalization of linear regression models that allows a wide range of error distributions and is expressed in terms of the mean

function. Mixed Model is a linear model that has coefficients which may be fixed or random. The clinical data we use consists of 109 patients on hormone therapy monitored over a period of six years. The patients at the end of the study were classified as non-eligible (6), non-compliant (5), loss to follow up (5), inter current illness (2), adverse experience (2), and administrative / other reasons (4). The preceding 24 patients were excluded from the modeling study. The remaining, were classified as death (17), treatment failure (26), an end of study (42).
These remaining eighty five patients were randomly divided into two groups. Group 1 patients' data (43) is used for model development and the remaining patients' data (42) as Group 2 for prediction. In Group 1 we focus on those patients for whom the hormone therapy was completed (21 patients). This means that the patients successfully reached the end of the study. This focus was chosen to find the ideal model for IAS therapy. Thus we developed four models as described below in this section. These models are used to predict PSA for all the patients especially those for whom the therapy did not work in Group 1and all the patients in Group 2. The effectiveness of the models is evaluated by predicting the PSA for all the eighty-five intent-to-treat patients and comparing them with their actual PSA reading especially for patients who did not reach the end of study of the IAS therapy. This comparison helps us to determine what it would take for patients who did not reach the end of study such as death to have shown better results.

The Group 1 randomly obtained data for modelling is studied using exploratory data analysis for any relationship among the different available variables as preparation for analysis. Exploratory data analysis summarizes the basic characteristics of the variables mainly using graphical methods. This enables us to observe the high correlation between testosterone and PSA during the on-treatment period among almost all Group 1 patients. In order to emphasize the importance of this observation and its generalizability, the same correlation is examined and computed over

all the available 109 patient data. Due to missing data this computation shows, correlation for the reduced number of 95 patients as can be seen from the Figure 1 below. Pareto Chart is a bar graph with descending order of the values of the variable.

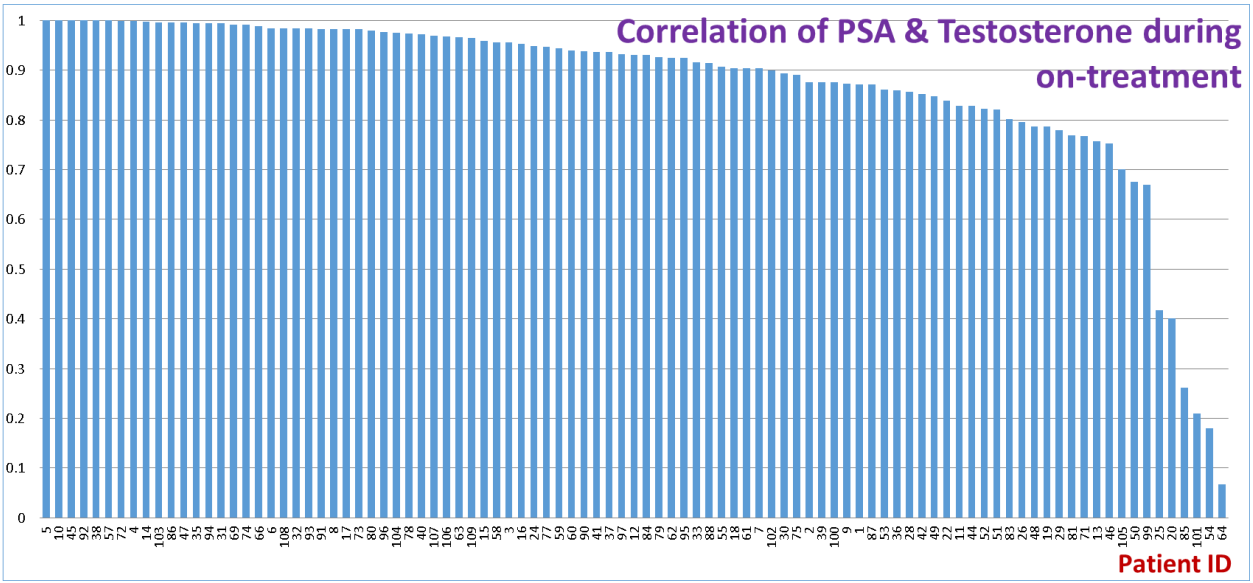

Figure 1. Pareto Chart showing positive linear relationship between PSA and Testosterone during IAS for on-treatment period.

As is evident from the graph, much more than majority of the patients' correlation of PSA and Testosterone during on-treatment period is above 90%. This is useful because it demonstrates that during the on-treatment period testosterone behaves as a surrogate variable to PSA. For a given patient, one could use linear regression to predict his PSA reading from testosterone readings or vice versa, with fairly reasonable accuracy.

The data can be viewed as consisting of two parts: one is the data collected at the start of the study, i.e., baseline data, and the other is longitudinal data observed until the end of six year study. The variables in the longitudinal data consist mainly of the variables such as File ID (identifying the patient), Site, the geographical site of the clinic where the patients visited, Date

of the visit, Week of the visit, PSA, Testosterone, Period, which is either 1 – on treatment, or 0 – off treatment and Cycle – the number of consecutive on- plus off-treatment periods the patient had gone through. The models are developed based on data from those patients in Group1 who successfully completed IAS treatment because we were searching for the "perfect PSA model" which would capture the behavior of the PSA in terms of the longitudinal covariates (i.e., independent variables that explain the variability in the dependent variable PSA). This focus was achieved by excluding in the perfect PSA model search, those patients who did not complete the therapy or those who died. Thus, our focus of developing the perfect PSA model was restricted to 21 patients from within Group 1 – the modeling group for whom the intermittent therapy was successfully completed. Since baseline data was observed only once there is no variability in these variables over time and hence they do not show up in the repeated measure models. Repeated measure models take into account the measurements which change with time. Also chemistry and hematological data was excluded from the search of the perfect PSA model.

The backward procedure which consists of including all the variables in the model and then eliminating the ones that are least significant based on p-values was used to arrive at the models. To avoid the problem of multicollinearity between covariates, we included only Date but not Week in the model. The data consisted of 1,522 records of which 1,116 are used but 406 are excluded due to missing observations. The error distribution was taken to be normal. Since some PSA readings would suddenly be very large and could also go as low as zero for a patient, in order that the error distribution could be taken to be normal, a natural log transform of 1+PSA is used. Thus the backward procedure using 'proc genmod' code within SAS program, with File ID as the repeated measure, and unstructured covariance matrix specification, yielded convergence of the algorithm for computing parameters with Generalized Estimating Equation

(GEE). Fit Criteria reading of QIC (Quasi-likelihood under the Independence model Criterion, Pan, 2001) = 1115.0862 and QICu = 1142.0000, the latter in addition penalizes in terms of the number of parameters used in the model (there is penalty for incorporating more parameters). These fit criteria were also used to get the best model by choosing the ones that have low fit values. The Table 1 below summarizes the results of testing the null hypothesis that the regression coefficient parameter of the model is zero against the alternative that it is not zero, from SAS output. The parameter estimate is the value of the function of the data estimating the regression coefficient of the Generalized Linear Model. The estimated standard deviations of these estimators are the standard errors (S.E.). The Z-statistics values are the estimate of the parameters divided by the corresponding S.E., which is therefore not specified in the Table 1 below, but give rise to the two-sided p-values from the standard normal distribution. The Generalized Linear Model in Table 1 was fitted with weights equal to the reciprocal of the sample variance. This model is identified from here onwards in this paper as Model 1. Note that by default SAS analysis binary variable in the numerical order, i.e., Period is taken by SAS program as an indicator variable of off-treatment because it is represented by "0". The summary of the results from SAS is as follows:

Table 1: Analysis Of GEE Parameter Estimates (Model 1)

Parameter	Parameter Estimate	Standard Error	P-value
Intercept	-16.8731	1.4758	<.0001
File ID	1.0024	0.1488	<.0001
= 14	0.2373	0.0774	0.0022
= 15	1.7514	0.1329	<.0001
= 22	0.4611	0.0346	<.0001
= 28	1.0103	0.1099	<.0001
= 31	0.7926	0.0290	<.0001
= 37	1.2891	0.1131	<.0001
= 40	0.7517	0.0326	<.0001
= 50	1.6079	0.1022	<.0001
= 58	0.6035	0.1133	<.0001
= 60	0.8143	0.1237	<.0001
= 61	0.4651	0.0531	<.0001
= 62	0.4945	0.0916	<.0001
= 77	0.8321	0.1119	<.0001
= 79	1.3575	0.0857	<.0001
= 87	1.2623	0.0970	<.0001
= 95	-0.4218	0.0161	<.0001
= 96	0.1008	0.0280	0.0003
= 97	0.3474	0.0214	<.0001
= 104	0.1969	0.0290	<.0001
= 106	0.0000	0.0000	.
Testosterone	0.1693	0.0133	<.0001
Period = 0	-0.0459	0.0394	0.2448
Period = 1	0.0000	0.0000	.
Testosterone•(Period = 0)	-0.0913	0.0113	<.0001
Testosterone•(Period = 1)	0.0000	0.0000	.
Date	0.0013	0.0001	<.0001
Cycle	-0.7891	0.0920	<.0001

All the covariates tried in the model are File ID, Date, Testosterone, Cycle, Period, including the

interaction term between Testosterone and Period, turned out to be significant. Note that although Period by itself did not show significant, however, as was expected from the Pareto Chart Figure 1 the interaction between Testosterone and Period is significant and hence all the corresponding main effect covariates are also retained in the model. The final Model 1 is then given by the linear combination of I{File_ID = i} with their corresponding parameter estimates along with the linear combination of Testosterone, one minus Period (1-Period), Testosterone multiplied by (1-Period), Date, and Cycle. The residual analysis shows that the histogram, boxplot, and normal probability plot of residuals are all indicative of the normal error distribution is the valid assumption. Furthermore, for the residual data, the Kolmogorov-Smirnov test statistics value D = 0.026633, yielding a p-value of 0.0532 > 0.05. All the graphs and this test indicate that the normal error with repeated measure design gives a good fit and especially so because the GEE parameter estimation method converged. All four models mentioned in this paper yield similar residual graphs as previously discussed. This model, which decides the inference of the parameters based on deviance analysis, is what we term as the first model (Model 1). Similarly, a model with the same weights as before was developed with similar answers using the SAS procedure of 'proc mixed', which develops the estimation using the Maximum Likelihood Method and the analysis of variance with covariance structure for the repeated variable File ID taken to be AR(1) (autoregressive with one parameter), with the following results in Table 2 to 5. This model from here onwards will be referred to as Model 3 in this paper, which is similar in results to Model 1. The Z-values in Table 2 are as defined in Table 1 and are therefore suppressed.

Table 2. Covariance Parameter Estimates (Model 3)

Covariance Parameter	Subject	Parameter Estimate	Standard Error	P-values
AR(1)	**File ID**	0.8787	0.01499	<.0001
Residual		0.02214	0.002677	<.0001

The following fit statistics were used to come up with the best Mixed Model. None of the random coefficient terms turned out to be significant. However Model 3 does confirm the results of Model 1 above because they give rise to very similar equations.

Table 3. Fit Statistics (Model 3)

-2 Log Likelihood	698.7
AIC (smaller is better)	754.7
AICC (smaller is better)	756.2
BIC (smaller is better)	698.7

In Table 4 below, the t – values similar to the Z-values are defined to be the parameter estimate divided by S.E. with degrees of freedom (D.F.) equal to 1090, except when estimating the intercept which has zero degrees of freedom and hence no p-value associated with it.

Table 4. Likelihood Solution to the Parameters of Fixed Effects (Model 3)

Effect of the Covariate	Parameter	Standard Error	P-value
Intercept	-12.9891	1.5915	.
File_ID	1.2965	0.5625	0.0214
= 14	0.1474	0.3900	0.7056
= 15	1.6463	0.4748	0.0005
= 22	0.2858	0.3604	0.4280
= 28	1.2204	0.5953	0.0406
= 31	0.4906	0.4539	0.2800
= 37	1.4391	0.4876	0.0032
= 40	0.3650	0.3632	0.3151
= 50	1.6996	0.5665	0.0028
= 58	0.6617	0.4354	0.1289
= 60	1.2435	0.5546	0.0252
= 61	0.4188	0.3960	0.2905
= 62	0.7849	0.6284	0.2119
= 77	0.8136	0.5218	0.1192
= 79	1.4010	0.4182	0.0008
= 87	1.5195	0.6358	0.0170
= 95	-0.1619	1.2573	0.8976
= 96	-0.02859	0.3517	0.9352
= 97	0.09776	0.3841	0.7991
= 104	0.2776	0.4683	0.5534
Testosterone	0.1306	0.003837	<.0001
Period = 0	0.01458	0.04122	0.7235
Period = 1	0	.	.
Testosterone•Period = 0	-0.09184	0.005274	<.0001
Testosterone•Period = 1	0	.	.
Date	0.001052	0.000113	<.0001
Cycle	-0.9285	0.07297	<.0001

Although some of the p-values corresponding to File ID in Table 4 do not show up to be significant, however, in the Type III analysis of variance results in Table 5 below shows File ID to be significant. The Mixed Effects Model can show random coefficients (as opposed to fixed effects coefficients shown in Table 5 below) when significant.

Table 5. Type 3 Tests of Fixed Effects Parameters (Model 3).

Effect of Covariate	Numerator D.F.	Denominator D.F.	F Value	P-value
File_ID	20	1090	3.16	<.0001
Testosterone	1	1090	903.89	<.0001
Period	1	1090	0.13	0.7235
Testosterone*Period	1	1090	303.25	<.0001
Date	1	1090	87.43	<.0001
Cycle	1	1090	161.93	<.0001

Thus Model 3, whose fit converged, gives similar results to the previous model and these models confirm each other's results because they give rise to very similar equations. The other models developed are those that do not use the File ID as a repeated measure but include all the meaningful baseline variables in a backward procedure to obtain two models (Model 2 and Model 4), one using 'proc genmod' and the other using 'proc mixed' through SAS, respectively. Similarly as before, both models with normal errors give analogous results to each other. In fact Model 4 residual analysis showed similar residual plots with only two mild outliers but the Kolmogorov-Smirnov test statistic D = 0.028869, p-value > 0.1500, Cramer-von Mises statistics W-Sq = 0.055751, with p-value > 0.2500 and Anderson-Darling statics A-Sq = 0.453359, with p-value > 0.2500, which are obviously not significant at the 5% level (i.e., normality of the residuals cannot be rejected). All the graphs and the tests indicate that the normal error for the

error distribution gives a good model fit. The validity of Model 4 is also confirmed because the algorithm to find Maximum Likelihood Estimates converged.

Although many meaningful and useful baseline variables related to prostate cancer study were taken into account, many of these variables had either missing data or did not show much variability in the data and therefore did not show up significant in the model. However, baseline variables such as registered age (Reg_Age), Body Mass Index (BMI), Blood Pressure (BP), Gleason Score (GleasonS), Radiation therapy (Rad_t), Total Prostate Volume (TPV) and number of Positive Cores (PosCores) showed up to be significant in these models to explain the variability in natural log of PSA+1 in addition to Testosterone, Period, Testosterone interaction with Period, Date and Cycle. Also, of the 1522 records of successful therapy patients in Group 1, only 481 records were used to develop these models (Model 2 and Model 4) due to missing observations in the baseline variables.

All four models are used to analyze the data from patient # 62 as an example. One could have used any other patient for whom the model fit is close. The results are as follows.

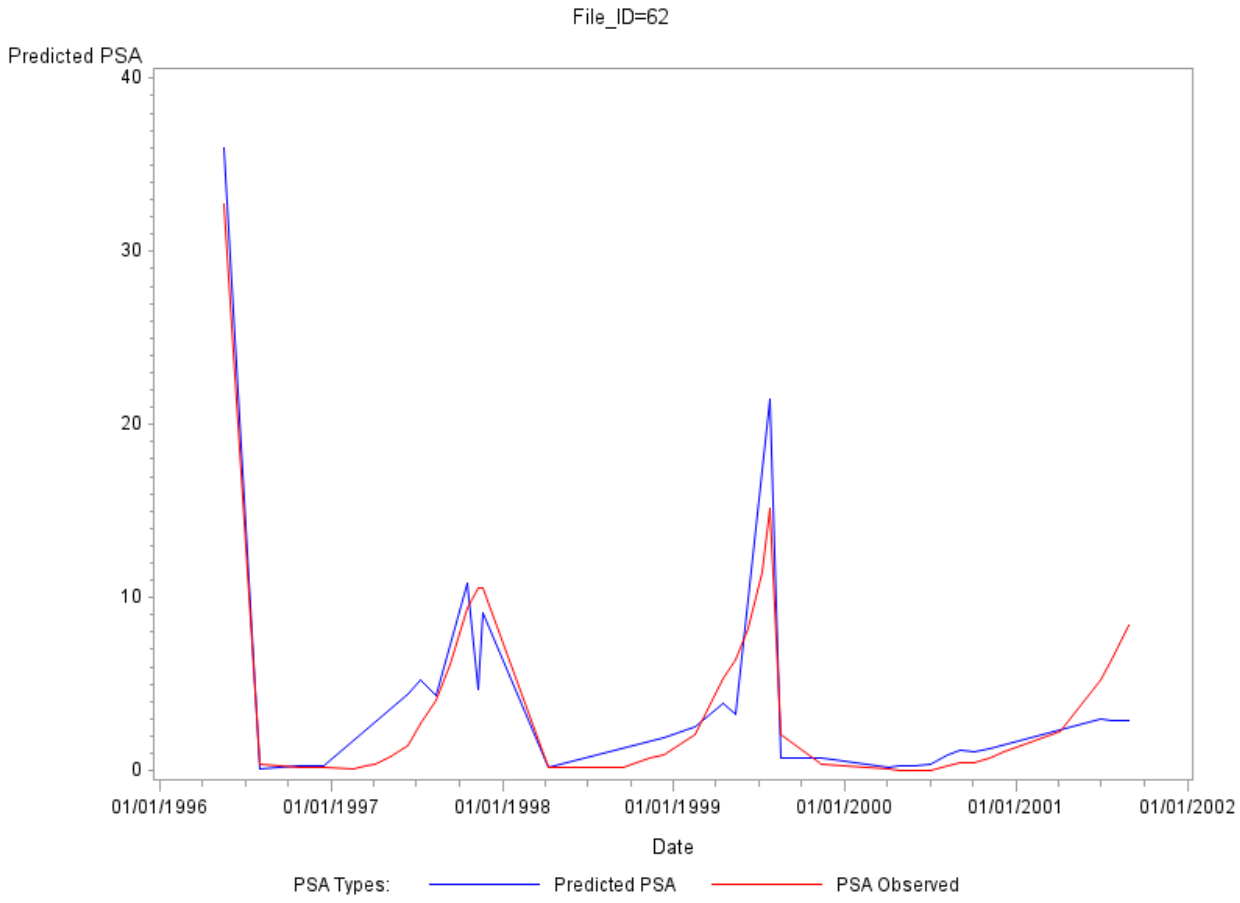

Figure 2. Plot of PSA and Predicted PSA versus Date (Model 1).

The PSA readings as a function of time during any given cycle form an elbow or V – shape, the left line of the "V" being formed by PSA readings during on-treatment period and the right line of the V by that of the off-treatment period. The **Model 1** equation is described by:

$$1+\text{PSA}(t) = f\left[X_1, X_2(t), X_3(t), X_4(t), X_5(t), t\right],$$

where X_1 = is the linear function of File_ID's, $X_2(t)$ = Testosterone, $X_3(t)$ = 1-Period, $X_4(t)$ = Testosterone multiplied by (1-Period), and $X_5(t)$ = Cycle, and t, which represents time, i.e., the variable Date, respectively. This function $f\left[X_1, X_2(t), X_3(t), X_4(t), X_5(t), t\right]$ also known as the

conditional expectation of response is

$$= \exp\{-16.8731 + X_1 + 0.1693X_2(t) - 0.0459X_3(t) - 0.0913X_4(t) - 0.7891X_5(t) + 0.0013t\}.$$

Here,

$$\begin{aligned} X_1 = {} & 1.0024I_{\{\text{File\_ID=13}\}} + 0.2373I_{\{\text{File\_ID=14}\}} + 1.7514I_{\{\text{File\_ID=15}\}} + 0.4611I_{\{\text{File\_ID=22}\}} \\ & + 1.0103I_{\{\text{File\_ID=28}\}} + 0.7926I_{\{\text{File\_ID=31}\}} + 1.2891I_{\{\text{File\_ID=37}\}} + 0.7517I_{\{\text{File\_ID=40}\}} \\ & + 1.6079I_{\{\text{File\_ID=50}\}} + 0.6079I_{\{\text{File\_ID=58}\}} + 0.8143I_{\{\text{File\_ID=60}\}} + 0.4651I_{\{\text{File\_ID=61}\}} \\ & + 0.4945I_{\{\text{File\_ID=62}\}} + 0.8321I_{\{\text{File\_ID=77}\}} + 1.3575I_{\{\text{File\_ID=79}\}} + 1.2623I_{\{\text{File\_ID=87}\}} \\ & - 0.4218I_{\{\text{File\_ID=95}\}} + 0.1008I_{\{\text{File\_ID=96}\}} + 0.3474I_{\{\text{File\_ID=97}\}} + 0.1969I_{\{\text{File\_ID=104}\}}. \end{aligned}$$

Note that $I_A = 1,$ if event *A* is true and 0, otherwise. A time t_1 is taken to be a time point between 17-May-96 to 2-Aug-96 taken from the left section of the V –shaped PSA graph, when Period is 1, i.e., on-treatment at time t_1, (1+PSA)(t) is written as (1+PSA)ON(t_1). Also, PSA readings during on-treatment are represented as PSA-ON. Similarly, when Period is taken to be 0, i.e., off-treatment at time t_2 it is (1+PSA)OFF(t_2), where t_2 is taken to be a time point between 13-Jun-97, 10-Nov-97, taken from right section of the V-shaped PSA graph. Also, PSA readings during off-treatment are represented as PSA-OFF.

Note that, since PSA during on-treatment is generally declining as a function of time and that PSA during off-treatment is inclining with a positive slope. Optimal scheduling for a fixed cycle is taken to be that period under which PSA-OFF ≤ PSA-ON. Optimal period is reached when PSA-ON = PSA-OFF because we do not want the patient to be too long on- off-treatment. If the patient is too long on- off-treatment then, their PSA will rise, however, we do not want them to be on- off-treatment period for too short of a time. Shorter off-treatment period translates in the next cycle into longer on-treatment which in turn would imply longer lower quality of life.

To achieve the preceding optimal condition, we consider the following quotient to be equated to '1':

(1+PSA)ON(t_1) / (1+PSA)OFF(t_2) =

Exp { 0.0459 + 0.1693 • (Test-ON – Test-OFF) + 0.0913 • Test-OFF + 0.0013 • ($t_1 – t_2$)},

where Test-ON and Test-OFF are testosterone variable during on- and off-treatment periods, respectively. The largest empirical duration of days between t_2 and t_1 computes to be 543 days using the ranges of times specified above.

Let PSA-ON = PSA-OFF, which implies

$$N - 0.0013 \bullet (D) = 0, \quad (1)$$

where $D = t_2 – t_1$ and N = 0.0459 + 0.1693 •Test-ON - 0.078 • Test-OFF. Thus D = N/0.0013.

Test-ON = 10.75 is an average reading in the range of t_1 specified above during Cycle 1, for patient 62. Similarly, Test-OFF = 17.26 an average reading in the range of t_2 specified above. Thus N = 0. 519595. Therefore, from equation (1) D = N/0.0013 = 400 days. The same procedure as described above for Cycle 1 can be followed within each of the Cycles 2 and 3.

Model 2 equation is given by $f\left[X_1(t), X_2(t), X_3(t), X_4(t), X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, t\right]$ similar to the one described in Model 1. Here, *f* is 1+PSA as a function of Testosterone, 1 - Period, Testosterone multiplied by (1-Period), Cycle, which itself are functions of *t*, i.e., $X_1(t), X_2(t), X_3(t), X_4(t)$, respectively. Also, $X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}$, represent the baseline variables registered age Reg_Age, body mass index BMI, blood pressure BP, Gleason score GleasonS, duration of radiation therapy Rad_t , total prostrate volume TPV, and positive cores PosCores, respectively. Thus,

$$f\left[X_1(t), X_2(t), X_3(t), X_4(t), X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, t\right] =$$

$$\exp\left\{\begin{array}{l} -23.6633 + 0.1786X_1(t) - 0.2112X_2(t) - 0.0956X_3(t) - 0.8294X_4(t) + 0.0011t \\ +0.1198X_5 + 0.066X_6 - 0.0519X_7 + 0.2004X_8 + 0.0493X_9 + 0.0299X_{10} + 0.3669X_{11} \end{array}\right\}.$$

Following similar calculations as in Model 1 above,

(1+PSA)ON(t_1) / (1+PSA)OFF(t_2) = Exp { 0.2112 + 0.1786 • (Test-ON – Test-OFF) + 0.0956 • Test-OFF + 0.0011 • (t_1 –t_2)}.

Letting PSA-ON = PSA-OFF, implies

$$N - 0.0011 \bullet (D) = 0, \qquad (2)$$

where N= 0.2112 + 0.1786 •Test-ON - 0.083 • Test-OFF and D = $t_2 - t_1$. Thus D = N/0.0011. Again, Test-ON = 10.75 and Test-OFF = 17.26. Thus N = 0.69857. Therefore from equation (2) D = N/0.0011 = 635 days. The computations for Models 3 and 4 are similar to Models 1 and 2, respectively. The models 1-4 based duration calculations when PSA-ON is set equal to PSA-OFF are 400, 635, 685 and 541, respectively. The duration results for the four models and for Cycles 1, 2 and 3 are presented in Table 6.

Table 6. Durations in days

	Predicted Duration				Empirical Duration
Cycles\Models	1	2	3	4	
1	400	635	685	541	543
2	79	239	251	168	540
3	1099	1514	1138	1370	763

As an illustration for using this predicted duration, consider the predicted duration of say 400 days and suppose that the patient was put on treatment for 250 days by the end of which his PSA

value came down to a normal level. The patient has to be then put on- off-treatment for the remaining 150 days. Then the new cycle starts.

Now to see the predictive power of the fitted models, we consider patient with File ID 3. This patient died during the treatment period. The predicted PSA values based on the first two models for this patient versus actual PSA readings are shown in Figure 3 below. Results for the remaining two models will be similar.

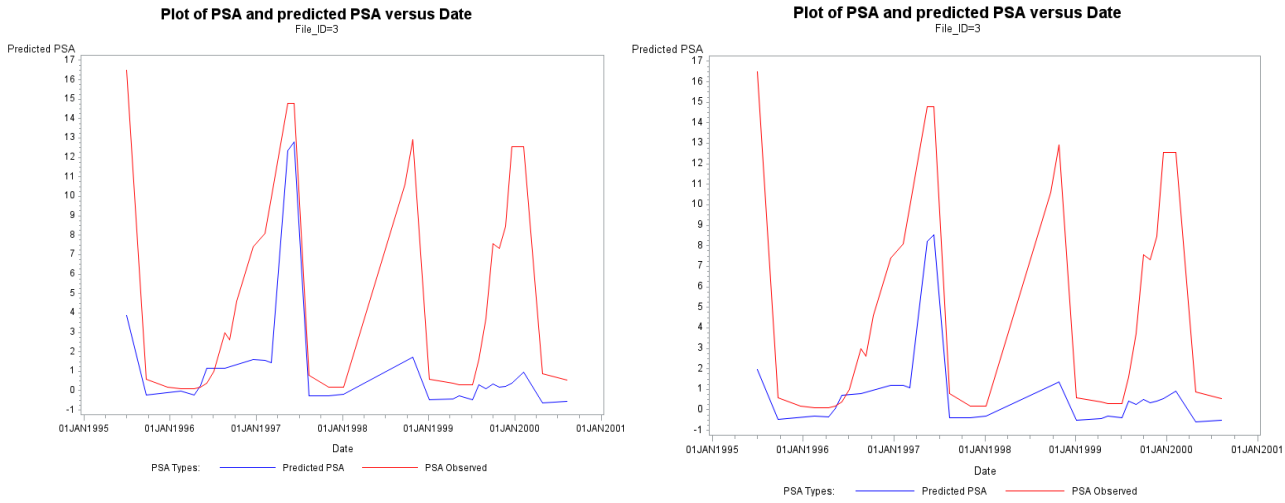

Figure 3. Based on a) Model 1 (left) and b) Model 2 (right).

We notice consistently that both models predict that the PSA reading for patient 3 should have been much lower across the course of the intermittent therapy, with very little overlap between the actual and the predicted PSA. Note that similar predictions are seen for all patients for whom the treatment failed or that died during the IAS six year therapy.

Discussion. In order to compare the finding of our research with those of Hirata et al. (2010-2, Section V, Numerical Examples and Figure 4), we have the following observations:

In their research three patient cases are considered. The third patient whose model based PSA is continuously rising, CAS is recommended without mentioning the therapy's impact on quality of life or the impact of other covariates such as testosterone on his monitored PSA. The first patient whose model based PSA is cyclically declining trend, an optimal schedule of first 259 days of on-treatment followed by 7 days of off-treatment is recommended, without regard to the patient's quality of life. Also, it is well known that PCa cells convert into AI cell over prolonged CAS. This treatment regime puts the patient on- off-treatment only for a week (compared to 259 days of on-treatment), which is almost a CAS therapy. Thus, on- and off-treatment periods' cycles of 266 days would continue. In case of the second patient whose model PSA is cyclically rising an optimal schedule cycles each of length of 721 days is recommended. An initial on-treatment period of only 168 days (about a half year), followed by 553 days off-treatment is considered here. It is not clear why such a period of more than a yearlong off-treatment is justifiable when the model clearly is predicting a growing trend in the patient's PSA. In view of the above observations no meaning comparison can be reported.

Conclusion. We observed in this study that a large majority of patients' correlations between PSA and Testosterone during the on-treatment period of IAS were at least 0.90. The length of time duration for a treatment cycle, while taking account of when the treatment should be "on" and when it should be "off," is calculated from a model, which is a useful application. To illustrate this, consider a patient under treatment. If his PSA level is low during on-treatment, when his therapy is stopped, then the number of days during the off-treatment period should be such that on-treatment days combined with off-treatment days (during treatment Cycle 1) should be no more than 541 days, according to Model 4 (or the largest empirical duration of 543 days). At this junction we would like the Oncologist to decide, how long the on-treatment period should

last depending upon the opinion of the patient in terms of his quality of life. Then, the duration of the off-treatment period can be determined by subtracting the on-treatment period from the 541 days of the total cycle time period. After 541 days, the PSA is likely to go up; therefore, the oncologist should put him on treatment until the PSA level comes down to the desired value in the next cycle.

Bibliography

Abrahamsson PA. (2010) Potential benefits of intermittent androgen suppression therapy in the treatment of prostate cancer: a systematic review of the literature. *Eur Urol.* 57, 41-59.

Akakura K, Bruchovsky N, et al. (1993) Effects of intermittent and, androgen suppression on androgen dependent tumor; apoptosis and serum- prostate antigen. *Cancer*. 71, 2782-90.

Bhindi, B, et al. (2015) Dissecting the association between metabolic syndrome and prostate cancer risk: analysis of a large clinical cohort. *European Urology* 67.1, 64-70.

Bruchovsky N, et al. (2006) Final results of the Canadian prospective phase II trial of intermittent androgen suppression for men in biomechanical recurrence after radiotherapy for locally advanced prostate cancer. *Cancer.* 107(2), 389-95.

Bruchovsky N, et al. (1990) Effects of androgen withdrawal on the stem cell composition of the Shionogi Carcinoma. *Cancer Res*. 50, 2275-82.

Bryant RJ, et al. (2015) Predicting high-grade cancer at ten-core prostate biopsy using four kallikrein markers measured in blood in the ProtecT study. *Journal of the National Cancer Institute* 107(7), djv095-1-6.

Chen Y, Hong C, Riley RD. (2015) An alternative pseudolikelihood method for multivariate random-effects meta-analysis. *Statistics in Medicine*. 34(3), 361-380.

Hirata Y, et al. (2012) Mathematically modeling and controlling prostate cancer under intermittent hormone therapy. *Asia Journal of Andrology*. 14, 270-277.

Hirata Y, et al. (2010-1) Development of a mathematical model that predicts the outcome of hormone therapy for prostate cancer. *Journal of Theoretical Biology.* 264, 517-527.

Hirata Y, et al. (2010-2) Hybrid optimal scheduling for intermittent androgen suppression of prostate cancer. *Chaos.* 20, 045125-1-10.

Hussain M, et al. (2013) Intermittent versus continuous androgen deprivation in prostate cancer. *N Engl J Med.* 368(14), 1314-25.

Moschini M, Sharma V, Zattoni F, Quevedo JF, Davis BJ, Kwon E, Karnes RJ. (2016) Natural History of Clinical Recurrence Patterns of Lymph Node-Positive Prostate Cancer after Radical Prostatectomy. *European Urolo*gy. 69(1), 135-42.

Mukherjee S. (2010) *The Emperor of all Maladies, a Biography of Cancer*. New York: Scribner.

Pan W. (2001) Akaike's information criterion in generalized estimating equations. *Biometrics.* 57, 120-125.

Tao R, Tavakoli M, Sloboda R, Usmani N. (2015) A comparison of US-versus MR-based 3-D prostate shapes using radial basis function interpolation and statistical shape models. *IEEE Journal of Biomedical and Health Informatics*. 19(2), 623-634.