VALIDATING RISK FACTORS FOR POSITIVE REPEAT TRANSRECTAL ULTRASOUND-GUIDED PROSTATE BIOPSIES USING DEMOGRAPHIC, CLINICAL, LABORATORY AND ULTRASOUND PARAMETERS

by

Yon J. Park

Presented to the Department of Public Health and Preventive Medicine

and the Oregon Health & Sciences University

School of Medicine

in partial fulfillment of

the requirements for the degree of

Master of Public Health

May 2002

School of Medicine
Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the M.P.H. thesis of

Yon J. Park

has been approved

[redacted]

Donald Austin, M.D., M.P.H., Committee Chair

[redacted]

Mark Garzotto, M.D., Member

[redacted]

Motomi Mori, Ph.D., Member

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The author wishes to thank Donald Austin, M.D., M.P.H., Mark Garzotto, M.D., and Motomi Mori, Ph.D., for their thoughtful help and contribution to this thesis. In addition, the author would like to thank the co-authors of the Portland Veterans Administration Medical Center Prostate Biopsy Database who include R. Guy Hudson, M.D., Eduardo Barrera M.D., Motomi Mori, Ph.D., Laura Peters, R.N., Thomas Klein, M.D., and Mark Garzotto, M.D. Finally, the author would like to thank Jennifer Keam, M.A., for reading drafts of this manuscript and providing feedback.

Abstract

Introduction - The purpose of the investigation is to determine risk factors for a positive repeat transrectal ultrasound (TRUS)-guided prostate biopsy in order to confirm the results of prior studies of men referred for repeat prostate biopsy. The aim of these studies was to identify risk factors of individuals initially negative for adenocarcinoma of the prostate (CaP) that place them at greater risk for a positive result upon subsequent biopsy. The rationale for conducting these studies is the variably high false negative percentage (10-34%) of the gold standard for diagnosis of CaP, the sextant TRUS-guided prostate needle biopsy. Discriminating characteristics between those with and without CaP have been determined in prior studies; however, there is no consensus about which risk factors are consistently associated with a positive repeat prostate biopsy. The purpose of this project is to confirm the findings of these prior studies using the Portland VAMC population.

Methods - The population is a prospective longitudinal cohort of 266 US Veterans referred for prostate evaluations and followed from February 1993 to July 2001. Each veteran was interviewed, examined and underwent at least one repeat sextant prostate biopsy. Data collected from each visit included age, race, referral indication, family history of CaP, prior vasectomy history, digital rectal examination (DRE), PSA indices, prostate volume, TRUS findings, initial biopsy result, number of biopsies, and follow-up period. Parameters significant

Abstract
(Continued)

in the univariate analyses, as well as variables found to be significant in prior studies, were analyzed in a multiple logistic model and a Cox proportional hazards analysis, respectively, in order to confirm findings of prior studies. Area under the receiver operator characteristics (ROC) curve analyses were conducted to assess the performance of the covariates in the models and to determine operating points, or "cut-offs."

Results - Sixty of 266 (22.6%) veterans had CaP on the final repeat prostate biopsy. Univariate analyses between subjects with benign and malignant final repeat prostate biopsies showed that initial PSA density ($p=0.000$), initial PSA ($p=0.02$), race ($p=0.034$), and referral indication ($p=0.044$) were statistically different. Multiple logistic regression analysis of variables found to be statistically significant to a $p<0.10$ showed that initial prostate volume ($p=0.008$), initial PSA ($p=0.01$) and initial DRE findings ($p=0.06$) were independent predictors of a final positive repeat prostate biopsy. Proportional hazards regression analysis of parameters found to be statistically significant in prior repeat prostate biopsy studies showed that initial PSAD ($p<0.00$), initial prostate volume ($p<0.02$) and the time-dependent variable number of prostate biopsy sets ($p=0.12$) were predictors of a final positive repeat prostate biopsy. Area under the ROC curve of initial PSAD interacting with initial prostate volume was 0.693 ± 0.04 ($p<0.00$).

Summary and Conclusions - PSA density, prostate volume, and DRE were confirmed as risk factors previously found to be associated with a positive repeat prostate biopsy. PSA velocity was not found to be a statistically significant predictor in this cohort. Cox proportional hazards modeling is a valuable tool in the assessment of risk for a positive final repeat prostate biopsy.

Introduction

Prostate Cancer is an important disease

The importance of prostate cancer for American men cannot be overstated. Prostate cancer remains the most incident cancer with the second highest cancer mortality (behind lung cancer) across all five racial and ethnic populations defined by the National Cancer Institute (NCI). After the spike of prostate cancer incidence in the early 1990s with the surge in popularity of PSA testing, the overall trend has been a slow increase in incidence with a decrease in mortality[1]. Most recently, the age-adjusted annual incidence is 144.6 per 100,000 for whites and 234.2 per 100,000 for Blacks, representing a +0.3% annual percent change and a +0.4% from 1995 to 1998, respectively.

Death rates are 22.4 per 100,000 for whites (-3.7% 1992-1998), 53.1 per 100,000 for blacks (-2.3% 1992-1998)[1]. In Oregon, death rates are 23.7 and 53.3 per 100,000 for White and Black men, respectively (1994-1998). Overall, prostate cancer accounted for 14.8% of cancer cases and 5.9% of cancer deaths[1]. In addition to the substantial incidence of prostate cancer, the prevalence of CaP has been estimated to be 30-40% of men older than 50 years of age[2].

These incidence and mortality data demonstrate the apparent contradiction of diagnosing and treating CaP: more will die with rather than from CaP. Dugan et al. (1996) defined the "clinically insignificant" cancer using mathematical modeling to project cancer volume doubling times relative to life expectancy tables. From 337 prostates examined, they defined clinically insignificant CaP as a "cancer that would grow to no larger than 20 cm^3 in

volume by the time of expected patient death and whose Gleason score was less than the tens place of the patient's chronological age.[2]"

The Gleason grading system is a widely used method for determining clinical significance of CaP: each tumor is graded by a primary (most commonly observed) and secondary (second most commonly observed) grade from 1 to 5. These grades are then summed as the Gleason score: well-differentiated tumors have a score of 2-4, moderately differentiated 5-7, and 8-10 scores are poorly differentiated[3]. Stamey et al. (1999) expanded on these criteria in a more recently published study of 379 prostates using a Cox proportional hazards model: they found that biochemical failure after surgical treatment was associated with the percentage of the cancer with Gleason grade 4 or 5 (Stanford modified Gleason scale), cancer volume, positive lymph node findings and intraprostatic vascular invasion[4]. Biochemical failure was defined as a rise in prostate-specific antigen (PSA).

Although prostate biopsy core Gleason scores are correlated with final Gleason scores (R^2 = 0.63[5], R2 = 0.57[4]), no algorithm exists to predict cancer volume preoperatively. Despite the apparent oxymoron of treating "benign cancer", data comparing the incidence and mortality in the US and the UK show that the rate of decrease in mortality is greater in the US, which the investigators suggest may be secondary to the more aggressive screening practices in the US[6].

Diagnosing Adenocarcinoma of the Prostate Gland is Difficult

Integral to the screening and diagnosis of prostate cancer, the systematic parasagittal sextant ultrasound-guided biopsy is considered the gold standard for obtaining specimens for pathological diagnosis of adenocarcinoma of the prostate (CaP)[7]. However, there is a high false negative ratio associated with the sextant prostate biopsy: in a study of 118 patients with known positive biopsies, Rabbani et al. (1998) found that 23% were negative on repeat biopsy[8] after a prior *positive* biopsy. Other studies have found false negative proportions ranging from 10% to 43% in repeat biopsy studies of men with prior negative biopsies[9-22]. The variability in proportion of false negatives is in part due to differences in study design. Such differences include performing repeat biopsies on the entire initially negative population (an exception rather than the rule) and using different prostate biopsying techniques (transperineal or saturation prostate biopsies).

Efforts made by the urologic community to address the proportion of high false negatives are reflected in the literature. There are many theorized causes for the high false negative proportion, the most obvious being random sampling error. Indeed, if each prostate biopsy core is approximately 0.002g[47], then a sextant biopsy represents only .06% of a 20 g prostate. Logically, the error increases as the prostate hypertrophies over time. This is corroborated by the observation that positive repeat biopsies are associated with smaller volumes[16] and that a negative prostate biopsy before a final positive prostate cancer diagnosis was associated with a heavier prostate gland[17].

Adenocarcinoma of the prostate is also multifocal in nature and the parasagittal sextant biopsy may not adequately sample crucial areas such as the anterior horn[24]. In order to increase the probability[25] many have increased the number of prostate biopsy cores per visit, thus increasing the sampling area and total volume of sample[18, 22, 24, 26, 27]. However, these efforts are limited by patient discomfort, increased morbidity, and criticisms that the cancers detected may be clinically insignificant. The increased sampling area of these biopsy schema also increases the likelihood of sampling smaller-volume tumors which have been associated with clinical insignificance[4].

The Findings of Repeat Prostate Biopsy Studies are Inconsistent

Finally, in order to improve the accuracy of the sextant prostate biopsy, many urologists have scheduled repeat prostate biopsy visits for men with a high suspicion for CaP who were negative on initial biopsy. However, what defines a "high suspicion" for CaP? Although many centers have undertaken improved classification of men at high risk using clinical, laboratory, and ultrasound parameters, a definitive and reproducible model has yet to be determined. A Medline literature search resulted in fourteen repeat prostate biopsy studies summarized in Table 1, representing a true variety of patient populations, study designs, analytical models and results. The univariate analyses include earlier studies as well as later studies involving greater than 6 prostate biopsies per visit.

PSA velocity (the annualized mean of at least 3 PSA values)[28] (Carter et al., 1992), was the independent variable found to be significant in the most

studies including Ukimura et al., 1997, Perachino et al., 1997, Borboroglu et al., 2000, although the cut-off varied from 0.75 ng/ml/yr[14] to 1.0 ng/ml/yr[13] . Ellis & Brawer (1995) found that there was a trend for PSA velocity to be significant in their case series of 100 subjects[10]. Another two studies found volume to be significant[16, 21], although no cut-off was published. Roehrborn et al. (1997) discovered the area under the receiver operating characteristics curve (ROC) for only absolute PSA value was statistically significant[11] while Letran et al. (1998) found that percent free-to-total PSA, the percent of free prostate-specific antigen to total free and bound prostate-specific antigen, was the only significant variable[15]. Finally, Stewart et al. (2001) found that only age was statistically significant in a saturation prostate biopsy study[22].

Table 1: Repeat Prostate Biopsy Studies

Studies	Subjects	Prevalence	Study Design	Analysis	Results	Comments
Keetch et al, 1994 WUStL	n=427	19% (82/427)	Prospective cohort	hierarchical multiple logistic regression	DRE (OR 3.5) Number of bxs (OR 0.2)	Interbiopsy interval significantly different for between groups (p = .0001)
Ellis & Brawer, 1995 Univ. of WA	n=96	20% (20/100)	Retrospective case series	non-paired t-test	PSA, PSA velocity not significantly different	Trend for PSA velocity to be significant
Roehrborn et al., 1996 UT Southwestern	n=123	23% (28/123)	Retrospective case series	univariate - t test, ROC	PSA ROC = 0.63 (p=0.02)	PSA velocity and PSAD unattainable
Fleshner et al., 1997 Sloan-Kettering	n=130	30% (39/130)	Prospective case series	multiple logistic regression	PSA > 20 (OR 4.48)	Mean f/u time for PSA velocity is 12.5 months
Perachino et al., 1997 Pietra Ligure, Italy	n=148	41% (60/148)	Prospective case series	univariate -chi-squared	mean PSA velocity (p=0.002)	PSA velocity > 1 ng/ml "high risk"
Ukimura et al., 1997 MD Anderson	n=193	26% (51/193)	Retrospective case series	univariate - chi-squared, McNemar's	ROC analysis: no sig. differences PSA et al.	PSA velocity > 0.75 ng/ml/yr: 13% vs. 42% p=.0007
Letran et al, 1998 University of WA	n=51	29% (15/57)	Retrospective case series	univariate - t-test	% free PSA (p=.05)	% free PSA cut-off 22%: 95% sens., 44% spec.
Rietbergen et al., 1998 The Netherlands	n=442	11% (49/442)	Prospective cohort	univariate - chi-squared, Mann-Whitney	Volume (p=.003)	PSA and PSA velocity were not significantly different (PSA velocity p=.083)
Epstein et al., 1999 Johns Hopkins	n=74	19% (74/395)	Retrospective case series	multiple logistic regression	Ant/lat location (p=.046) Weight (p=.002)	Prior benign bx study of RRP specimens. PSA velocity predicts tumor volume (p<.0001)
Borboroglu et al., 2000 San Diego Naval	n=57	30% (17/57)	Prospective case series	univariate - rank sum, Fisher's exact	PSA velocity 0.9 ng/ml*yr 91% sens., 77% spec.	Extensive repeat bx study; mean f/u time 10.5 months. % free PSA not significant
Djavan et al., 2000 Vienna/ Brussels	n=820	10% (83/820)	Prospective cohort	multiple logistic regression	TZ-PSAD (p=.001) %free PSA (p <.001);	%free PSA cut-off 30%: 90% sens., 50% spec.
Fowler et al., 2000 Mississippi VA	n=298	27% (80/298)	Prospective cohort	multiple logistic regression	Age (p=.002), % free PSA (p=.0003) significant	Univariate analyses: age, PSA, PSAD, PSA velocity, %free PSA significant
Igel et al., 2001 Mayo Clinic, FL	n=88	43% (38/88)	Prospective case series	univariate - chi-square, Wilcoxon	Volume (p < .001)	Transperineal bx study; 94% of pts were chosen for PSA velocity > 0.75
Stewart et al., 2001 Mayo Clinic, MN	n=224	34% (77/224)	Prospective case series	univariate - Wilcoxon	Age p=.02	Saturation bx study; interbiopsy interval unknown for PSA velocity

Keetch et al. (1994) were the first group to employ a multivariate model[9]. In their final analysis, only digital rectal examination (DRE) findings and the number of biopsies were statistically significant. None of the PSA indices tested were significant, although their analysis of PSA velocity was limited by a significant difference in inter-testing period between groups. PSA velocity is inversely proportional to time between testing[23] and a significant time difference can directly affect the calculation of this parameter. In other words, holding the change in PSA values equal, a shorter time period will have a greater PSA velocity than a longer time period. Epstein et al. (1999) found that PSA velocity predicted tumor volume in their retrospective case series of 74 patients, but were unable to find a clinically meaningful cut-off value[17]. They also found that an anterior or lateral location and prostate weight were predictive of a negative prior biopsy before radical retropubic prostatectomy (RRP).

Fleshner et al. (1997) found only PSA above 20 ng/ml to be predictive[12]. Both Djavan et al. (2000) and Fowler et al. (2000) found %-free-to-total PSA to be highly statistically significant in their multiple logistic models[19,20]. In addition to %-free-to-total PSA, Djavan et al. (2000) also found transition zone volume PSA density (TZ-PSAD), or the ratio of total PSA to the volume of the transition zone of the prostate, to be significant in their model. The transition zone is an area of the prostate immediately adjacent to the proximal prostatic urethra. In young adult men, this area accounts for approximately 5% of the total volume of the prostate[29]. Fowler et al. (2000) also found age to be significant in their model with %-free-to-total PSA[20] . Djavan et al. (2000) were limited in their analysis of

PSA velocity because of an inter-testing period of 6 weeks[19]. Because PSA velocity is an annualized calculation, an inter-testing time period less than 6 months would artificially elevate PSA velocity. This is a limitation that is expressly cautioned against by Carter et al. (1995), the group responsible for defining this clinical entity[23].

Two Prior Studies Closely Resemble the Cohort of this Investigation

The patient population to be studied in this investigation most closely resembles the groups in Roehrborn et al. (1996), and Fowler et al. (2000), and those studies will thus be examined closely. Roehrborn et al. sought to determine the diagnostic yield of repeat biopsies in a retrospective case series of 123 veterans referred to the Dallas Veterans Affairs Medical Center for repeat sextant prostate biopsy for an elevated PSA and/or an abnormal DRE. The published results of this study do not include information about demographic, clinical or ultrasound findings. Furthermore, only one laboratory variable, PSA, was included in the analyses. The Student's t-test was used to analyze the PSA value between those patients with and without CaP, and was not found to be statistically significant. However, the validity of the test for the data is questionable since PSA data do not follow a Gaussian distribution. As mentioned, the ROC for PSA was statistically significant at 0.63 ± 0.06 SE ($p=0.02$)[11].

Fowler et al. (2000) also conducted a study of US Veterans (Mississippi Veterans Affairs Medical Center): 298 black (55%) and white (45%) consecutive male veterans underwent a repeat prostate biopsy due to an abnormal DRE

and/or PSA >= 4.0 ng/ml after an initially negative prostate biopsy. Variables analyzed included age, race, DRE, biopsy technique, presence of prostatic intraepithelial neoplasia (PIN), the Gleason scores of positive biopsies, PSA, PSA density, PSA velocity and %-free-to-total PSA. PIN is considered to be the precursor to malignant cancer. After initial univariate analyses determined that age, PSA density, %-free-to-total PSA and PSA velocity were significant, a multiple logistic regression revealed that only age and PSAD (the ratio of serum prostate-specific antigen to prostate volume) were statistically significant. In subset analyses using patients in whom PSA velocities and %-free-to-total PSA results were available, age and %-free-to-total PSA were statistically significant.

Notably, Fowler et al. (2000) did not find PSA velocity or volume to be significant, but did replicate findings by Letran et al. (1998) and Djavan et al. (2000) with regards to %-free-to-total PSA. However, the validity of their PSA velocity determinations is questionable. In their discussion, Fowler et al. (2000) state, "PSA velocity determinations were derived in part from PSA obtained after the last biopsy and the data were not available in all patients at that time."[20] It is unclear in the methods section which of the subjects had PSA measurements gathered after the last biopsy. When using PSA values after the last biopsy, it cannot be known if CaP had developed in the prostate in the interim after the last biopsy and last recorded PSA. Therefore, high PSA velocities may actually be representative of a positive subject misclassified as a negative.

Also, readers must assume that PSA measurements would not be recorded in those with positive biopsies who had already received therapy

because medical or surgical treatment for CaP markedly reduces PSA values. Finally, assuming the authors only recorded PSA values ex post facto for those with negative last biopsies, then a bias may exist towards recording lower PSAs since men without CaP have lower PSAs[30]. If many of their data used to calculate PSA velocity were obtained from negative subjects, then PSA velocity calculations would be lowered artificially. Appropriately, they state the role of PSA velocity for assessing risk for positive biopsy as inferred from their study is "unclear."[20]

Confirmation of Prior Repeat Biopsy Studies is Necessary

The purpose of this investigation is to develop a model that discriminates between patients positive and negative for CaP on repeat biopsy using demographic, clinical, laboratory, and ultrasound parameters. Also, in order to confirm the findings of previous studies attempting to identify those individuals with negative initial biopsies at greater risk for a positive repeat biopsy (please see Table 1), the results of prior studies will be incorporated into a proportional hazards model. These results may not be validated, but part of the process will involve using a different statistical model, which may be better suited to time-dependent data. The purpose of this investigation mirrors the many studies preceding it; however as the role of various clinical and laboratory indices remains unclear, the justification for this study remains evident.

Subjects, Apparatus, Procedure

Data Collection

The population under investigation is a prospective longitudinal cohort of US Veterans referred to the Portland Veterans Administration Medical Center (PVAMC) for prostate evaluation and followed from February 1993 to July 2001. The population is a consecutive cohort exposed to referral for prostate examination by a urologist with no sample taken within this population. Information for this study was initially collected before human subjects regulations were formalized at PVAMC and subsequent data collection was informally approved. Institutional Review Board approval has since been received.

Each of 266 veterans were interviewed, examined and underwent one or more repeat sextant prostate biopsies. Exclusion criteria included prostate biopsies that were not at least sextant and a previous diagnosis of CaP (in 15 patients, a prior transurethral resection of the prostate (TURP) had diagnosed CaP, or the repeat biopsy was used to monitor progression of CaP). Types of data collected are tabulated in Table 2 and include age, race, referral indication, family history, prior vasectomy, DRE findings, PSA, PSA density, rate of change of PSAD, PSA velocity, PSA doubling time, prostate volume, TRUS findings, presence of prostatic intra-epithelial neoplasia (PIN), number of biopsies, interbiopsy period and follow-up period.

Table 2: Data collected for the PVAMC cohort

Category	Variables
Demographic	age, race
Clinical	referral indication, family history of CaP, vasectomy history, digital rectal exam finding
Ultrasound	hypoechoic lesions, volume
Laboratory/ Pathology	PSA, PSA velocity, PSA doubling time, PSAD, presence of PIN, number of biopsies, biopsy result
Time	follow-up, interbiopsy period

Race was coded as white, black, Hispanic, Asian/Pacific Islander, and American Indian. Because this study is based entirely on the Portland VAMC, the cohort reflects the small proportions within ethnicities and large proportion of non-Hispanic white within Oregon. Referral indications included abnormal DRE and/or elevated PSA, bladder outlet obstruction symptoms, or prior PIN. Positive family history was defined as having a first degree relative with CaP. Prior vasectomy history was recorded as positive if the patient had ever undergone a vasectomy.

Digital rectal examination findings were confirmed by a urologist and were coded as positive if they were suspicious or likely for cancer. Asymmetry was not considered to suggest a suspicious lesion in this study. Transrectal ultrasound was performed by a urologist using a Bruel and Kjaer system 1846 or 3535 ultrasound machine with a biplanar Endosonic Transducer Type 8551 using 7 MHz for the 1846 type and 7.5 MHz for the 3535. Hypoechoic lesions were coded as positive; other lesions such as calcifications and/or isoechoic findings were not considered positive. These ultrasound data are not complete for all patients because of the absence of a urologist (n=40). The volume of each prostate was determined using the prolate ellipsoid formula (p/6 x length x width

x height)[31].

PSA Indices

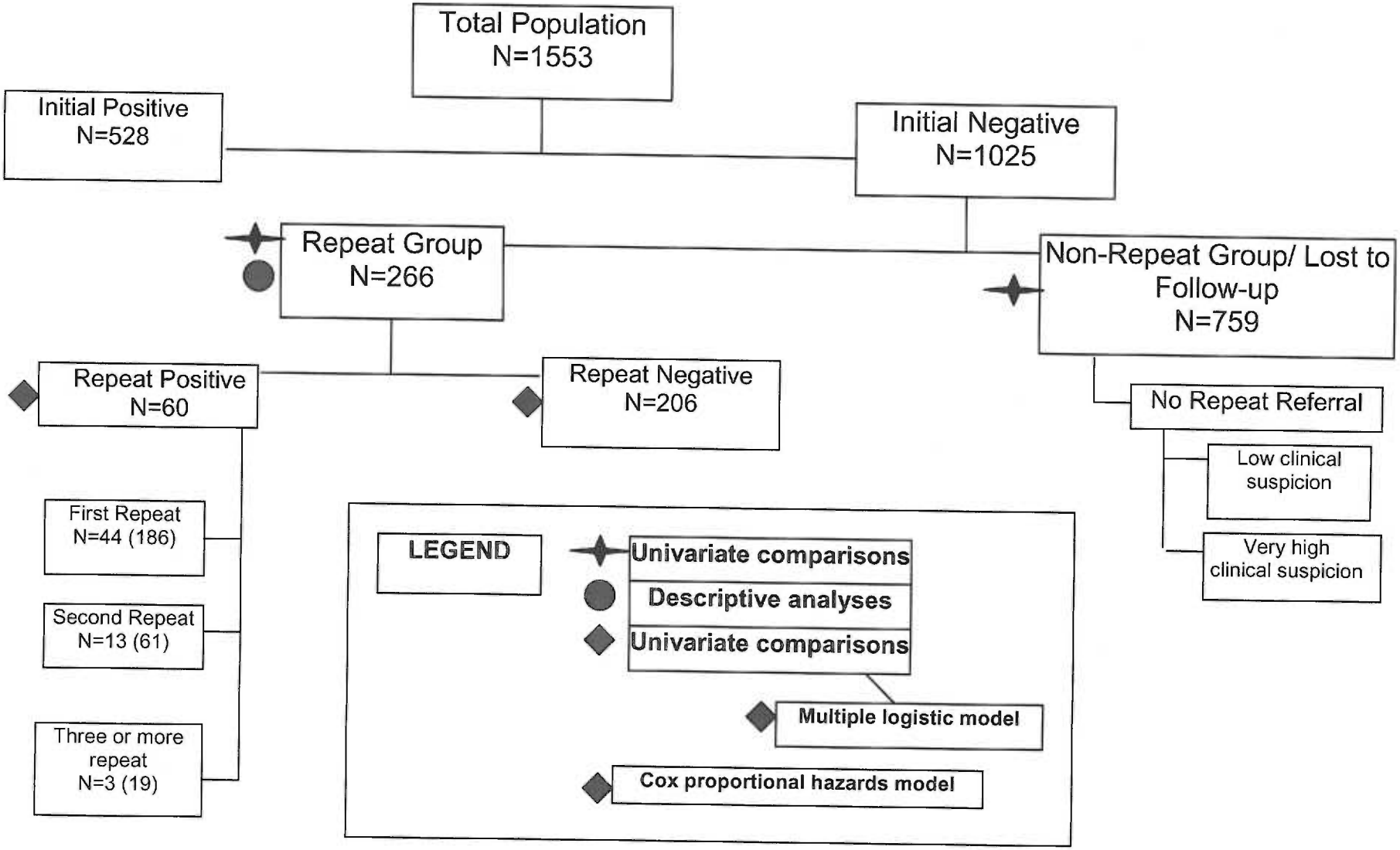
PSA was determined from sera collected at each biopsy visit before prostate perturbations and analyzed at the PVAMC laboratory using the Hybritech method. PSA density was calculated by dividing the PSA by the volume of the prostate at the time of the visit. PSA velocity was determined by determining the annualized mean of at least 3 PSA values collected before the last biopsy (n=168, 63.2%) using the method described by Carter et al. (1992): if multiple PSA measurements collected, the largest measurement closest to 6 months before the biopsy visit was chosen[18]. PSA doubling time was calculated using the formula described by Schmid et al., 1993: T= (log 2 x t)/(log PSA final-log PSA initial)[32]. Each of at least six biopsy core was collected using a Biopty™ (Bard Inc.) biopsy gun with an 18 gauge needle, in a distribution of the prostate gland developed by Terris et al., 1991[31], labeled with the patient's name, processed and microscopically examined by a pathologist. Follow-up period was defined as the time between the dates of the first recorded PSA measurement and the last biopsy.

Data Analyses

Descriptive analyses were performed on the demographic, clinical and laboratory data (Please see Figure 1, boxes marked with a circle), the results of which are in Table 2. Each continuous variable (age, PSA indices, prostate

volume, time) was plotted against a normal curve and tested for normality using one-way Kolmogorov-Smirnov statistics with a Lilliefors significance level for testing normality. Extreme values were also pursued and removed from analyses if they were artifactually elevated during calculation (as in PSA velocity or PSA doubling time).

Figure 1: Flowchart of Study Group Allocation



The population of veterans who underwent multiple biopsies was selected from the referred population for continued monitoring after initial negative sextant prostate biopsy. Of the 1553 subjects (please see Figure 1) who underwent sextant prostate biopsy, 1025 were initially negative, yet only 266 (26.0%) underwent repeat prostate biopsies. The factors influencing the selection of this small subpopulation are unclear, but many factors could have contributed. These may include loss to follow-up, the primary care provider not making a repeat referral, and the urologist not recommending repeat biopsies based on clinical suspicion. However, this population is an accurate representation of a clinical cohort seen by community and academic urologists who work with fluid and rapidly changing populations of patients and referring primary care providers. The question still remains for a urologist in paring down the large group of initially negative patients who raise clinical suspicion of false biopsy results. The aim of this study is to replicate and confirm the parameters previously used to develop that clinical suspicion.

Despite the clinical relevance of the cohort in this study, there is an acknowledged selection bias. In order to expose any possible biases in the data attributable to this recognized selection bias, univariate analyses were carried out to compare those selected, or self-selected, for repeat biopsy compared to those who did not return for repeat biopsy after initial negative biopsy. The variables analyzed were age, race, referral indication, family history, prior vasectomy, DRE findings, PSA, PSAD, prostate volume, TRUS findings, and presence of PIN. Statistically significant differences between the repeat and non-repeat groups,

especially in those clinical parameters already incorporated into clinical suspicion such as age, race, DRE findings, and PSA indices were expected. (Please see Figure 1, boxes marked with a star)

Although there is a recognized selection bias of the patients in this population, this selection bias is the rule rather than exception in the urologic literature. In the studies listed in Table 1, the only study designed to avoid this selection bias was that conducted by Djavan et al., 2000[19]. The prevalence of CaP in this cohort was determined as a basis for comparison to the other study patient populations of Table 1. Other demographic and clinical variables were also described, but the primary comparison used by urologists to gauge generalizability to their patient population is overall prevalence of CaP.

An estimation of incidence was calculated due to the highly variable time of follow-up in this patient population, which may be useful as a method of comparing patient populations in future studies. The prevalence was determined by dividing the number of patients with positive prostate biopsies by the total population with multiple biopsies. The 1, 2 and 5 year incidence of CaP was determined using Kaplan-Meier survival curves and the survival functions at these times. (Please see Figure 1, boxes marked with a circle)

Univariate analyses employed the Pearson chi-square test for categorical variables, the Mann-Whitney U test for variance for continuous variables with a non-normal distribution, and the Student's t-test for continuous variables with a normal distribution (please see Figure 1, boxes marked with a diamond). The variables found to be significant at an alpha of

0.10 (two-tailed) in the univariate analyses were considered as candidates for a multiple logistic regression model.

In an effort to uncover possible multicollinearities between variables, pairwise correlations using Pearson correlation coefficients for those variables with normally distributed data and Spearman's rank correlation coefficients for parameters with non-normal data were carried out. Pairwise correlation coefficients of moderate to large effect (Cohen, 1987) were incorporated into a multiple logistic regression model using an interaction term along with the other variables as a systematic approach to address possible multicollinearities. The more general approach of using correlation matrices for assessing the severity of multicollinearity was not employed.

The last biopsy result (either CaP or negative) was the dependent variable. The procedure used was a backward stepwise method using likelihood ratios and a Hosmer and Lemeshow goodness-of-fit index to assess overall model fit. Studentized residual plots were generated to search for the presence of outliers or leverage points. Cook's distance statistics were also generated to determine the influence of a data point on the regression equation[33]. The precedent for using univariate analyses and then a multiple logistic model is well-established[9, 12, 17, 19, 20] (please see Figure 1, boxes marked with a diamond).

The variables found to be statistically significant to an alpha of 0.05 (two-tailed) in the univariate and multivariate analyses in prior studies were incorporated into a Cox proportional hazards model[34] using this data set. Time was defined as the period between the first negative biopsy and the last biopsy.

If the last biopsy was positive, that event was defined as terminating, and if the last biopsy was negative, that subject was censored. The following variables were considered for incorporation into the model: DRE result, number of repeat biopsy visits, PSA, PSAD, PSA velocity, volume and age. A backwards stepwise procedure using maximum likelihood ratios was employed.

As in the development of the multiple logistic regression model the variables were tested before entrance into the model for possible multicollinearities. Pairwise Pearson correlation coefficients were used for those variables normally distributed data and pairwise Spearman's rank correlation coefficients were used for parameters with non-normal data. Variables with correlation coefficients of moderate to large effect, which may indicate severe multicollinearity, were incorporated into the model using interaction terms.

The Cox proportional hazards model incorporated time-dependent interaction terms when a comparison of survival curves of independent variables showed a violation of the proportional hazards assumption with a variable hazard ratio over time[33]. Another time-dependent variable was created for the clinical interaction of time and PSA velocity. The significance of this variable's coefficient was used to test if the proportional hazards assumption was reasonable. The regression model was tested for the ability to adequately describe the data with DfBeta statistics, which estimate the change in a coefficient if a case is removed[33].

The precedent for using a proportional hazards analysis in the urologic literature has been set by at least two groups, Katcher et al., 1997[35] and Stamey

et al., 1999[36]. The justification for its use in this study is an inherent characteristic of a multiple logistic model: its estimates are derived from data pooled or averaged over time, as opposed to taking time into consideration. In a clinical population, where follow-up periods are likely to be highly variable, the ability to consider time may be valuable. If time is a significant confounder, then a proportional hazards model will differ from a multiple logistic model[37]. Time is a confounder in the calculation of PSA velocity; absolute value is indirectly proportional to the amount of time elapsed. The highly variable follow-up periods in this population also may lead to a misclassification bias in a multiple logistic model, i.e., short-term follow-up patients classified as negative for CaP might have become positive had they been a long-term subject.

Area under the receiver operating characteristics curve (ROC) analyses were carried out on the variables found to be significant in the multiple logistic model and Cox proportional hazards model[38]. In addition to providing another measure of comparability with other studies, these analyses provide a means of establishing a "cut-off" value for variables found to be significant, but are continuous. The best operating point was chosen at a point on the ROC that lies at a 45-degree line closest to the (0,1) point of the ROC[39]. It is acknowledged that the results may be inflated because of the random error involved in recursively applying a test to a sample that created the model[40].

Although they were applied to the same dataset, the two different models are expected to have different results because of variable follow-up periods. Applying the standard multiple logistic regression model to this population may or

may not yield similar results as found in previous studies employing multivariate models. However, the Cox proportional hazards model, using variables already found significant in repeat prostate biopsies, may confirm the results of previous repeat prostate biopsy studies.

The two regression models were compared for ability to adhere to assumptions, for goodness-of-fit to the dataset, and for ability to describe the data adequately based on regression diagnostics. In the end, a definitive validation cannot be carried out because of the limitations in data collection (%free-PSA, TZ-PSAD), which will be discussed in the Discussion section. However, a potentially useful model in repeat prostate biopsy research will be introduced into the armamentarium of urologists.

Results

Characteristics of the Repeat Prostate Biopsy Group

There were 266 subjects who underwent one or more repeat prostate biopsies of a total group of n=1025 initially negative subjects (26.0%). Descriptive analyses were carried out on this population (please see Figure 1, boxes marked with a circle) and the results are summarized in Table 3. Briefly, the demographic data of the repeat prostate biopsy group indicate that the median age was 66 years, the race predominantly non-Hispanic White (91.3%) referred for elevated PSA and/or abnormal digital rectal examination (96.9%). Family history for CaP in a first-degree relative and prior vasectomy history was negative in the majority (81.2% and 67.2%, respectively). Clinically, a slim majority of veterans were normal for DRE (55.1%) and TRUS of the prostate (52.3%). However, 98 of 265 (37%) veterans had a suspicious or "cancer likely" DRE and 125 of 262 (47.7%) veterans had hypoechoic lesions on TRUS, two findings which have been associated with CaP in previous studies[9,41].

The median prostate volume was 35.75 cm^3 with an interquartile range (25^{th} to 75^{th} percentile) of 22.35 cm^3. PSA indices are summarized in Table 3; notably, PSA velocity, PSA doubling time and PSA density have wide interquartile ranges. Extreme values for PSA velocity included 4 cases at the high and low end (range -168.8 to 64.4): when plotted against time, these four cases are clearly artificially elevated due to the inverse relationship of time to PSA velocity (please see Appendix A: Figure 2). There was an outlier in the data for doubling time caused by an error in calculation by the software package. The

four data points of PSA velocity have been removed from analyses and recoded as missing data, and the PSA doubling time recalculated.

Table 3: Descriptive characteristics of the repeat prostate biopsy group

Variable	Subjects	Median	Interquartile Range	Percentage
Age at entrance (years)	266	66.00	8.25	
Race	264			
White				91.3%
Black				4.9%
Hispanic				2.3%
Other				1.5%
Referral Indication	265			
•PSA				66.0%
Abnl DRE				15.1%
Abnl DRE & •PSA				15.8%
Other				3.0%
Family History of CaP	266			
No				81.2%
Yes				18.8%
Vasectomy History	265			
No				67.2%
Yes				32.8%
Initial Digital Rectal Exam	265			
Normal				55.1%
Asymmetric				7.9%
Suspicious/ Cancer likely				37.0%
Initial TRUS findings	262			
No hypoechoic lesions				52.3%
Hypoechoic				47.7%
Initial Volume (cm^3)	266	35.75	22.35	
Initial PSA (ng/ml)	266	6.40	4.23	
PSA velocity (ng/ml)	168	0.56	2.28	
PSA doubling time (years)	247	6.14	9.35	
Initial PSAD ($ng/ml/cm^3$)	266	0.16	0.13	
Presence of PIN (ever)	263			
No				80.6%
Yes				19.4%
Number of repeat prostate biopsies	266	2.00 (mode)		
2				69.9%
3				22.9%
4 or more				7.1%
Gleason score >= 7	57			
No				66.7%
Yes				33.3%
Follow-up time (months)	266	18.23	23.60	
Interbiopsy time (months)	266	16.18	26.41	

The majority of subjects had just one repeat biopsy set (69.9%), but 80 subjects had at least a second repeat biopsy (38.8%, 80/206) and another 19 subjects had three or more biopsies (9.8%, 19/193). Fifty-one veterans (19.4%) had presence of prostatic intra-epithelial neoplasia, a pathological finding variably associated with CaP or progression into CaP[10]. Of those subjects diagnosed with CaP, nineteen veterans (33.3%) had evidence of clinically significant disease as demonstrated by a Gleason score of 7 or greater, a score associated with a worse prognosis for CaP[17,42].

In addition to these data and statistics, tests for normality were performed using one-way Kolmogorov-Smirnov statistics with a Lilliefors significance level. There were no parameters with data that followed a Gaussian distribution, and the statistical tests that follow will reflect this limitation.

The Non-repeat Prostate Biopsy Group Differed from the Repeat Group

To corroborate the predicted selection bias, there were statistically significant differences (two-tailed, $p < 0.05$) between those patients who underwent repeat prostate biopsies and those who did not (please see Table 4). In Table 4, "Other" includes lower urinary tract symptoms and/or PIN. Some of these differences are clearly related: more in the repeat prostate biopsy group were referred for elevated PSA and as a result, the median PSA and PSAD are likewise higher in the repeat group. In summary, those in the repeat prostate biopsy group were younger, were more often referred for elevated PSA, were more likely to have a normal DRE, were less likely to have PIN and had larger

prostate glands.

The repeat group is younger, a paradoxical observation in light of the increasing incidence of CaP with age[1]. Another incongruous finding is that the non-repeat group were much more likely to have an abnormal DRE. These differences in the two groups may reflect the difficulty of determining who will die from versus with CaP. A younger, healthier subject may have self-selected for repeat prostate biopsy, or may have been referred by a primary care provider or urologist for repeat prostate biopsy due to greater survival after diagnosis and treatment of CaP. The corollary is that an older subject with a positive DRE would not be referred for a repeat biopsy because of the lower disease-specific mortality of CaP[43]. Conjecture aside, it is unknown how the non-repeat prostate biopsy group differed from the repeat prostate biopsy group in overall prevalence of CaP because these data were not collected.

Table 4: Significant differences between non-repeat and repeat groups

Variable	P	Non-repeat (N=756)	Repeat (N=266)
Age at entrance (median, years)	0.007	68	66
Referral Indication	0.012		
Elevated PSA		53%	66.0%
Abnl DRE		24.6%	15.1%
Abnl DRE & +PSA		16.0%	15.8%
Other		4.3%	1.1%
Initial Digital Rectal Exam	0.000		
Normal		43.7%	55.1%
Suspicious/ Cancer likely		51.2%	37.1%
Asymmetric		5.0%	7.9%
Presence of PIN (ever)			
No		92.2%	87.9%
Yes		8.0%	12.1%
Initial Volume (median, cm^3)	0.048	33.5	35.8
Initial PSA (median, ng/ml)	0.047	5.9	6.4
Initial PSAD (median, $ng/ml/cm^3$)	0.025	0.14	0.16

Table 5 summarizes the non-statistically-significant differences, or the

similarities, between the repeat and non-repeat prostate biopsy groups. Briefly, race, family history of CaP, prior vasectomy history and TRUS findings were not statistically different between the repeat and non-repeat prostate biopsy groups. In Table 5, "Other" for race includes Asian/Pacific-Islander, American Indian and other. "Other" for referral indication includes lower urinary tract symptoms (GU sxs) alone, abnormal DRE & GU sxs, unknown, and other.

Table 5: Non-significant differences between non-repeat and repeat groups

Variable	P	Non-repeat (N=756)	Repeat(N=266)
Race	0.613		
White		91.3%	93.7%
Black		4.9%	3.6%
Hispanic		2.3%	1.0%
Other		1.5%	1.7%
Family History	0.431		
Yes		81.2%	84.2%
No		18.8%	16.3%
Vasectomy History	0.697		
Yes		67.2%	68.7%
No		32.8%	31.3%
Initial TRUS findings	0.301		
No hypoechoic lesions		52.3%	49.1%
Hypoechoic lesions		47.7%	50.9%

The Prevalence of CaP is Similar to Other Repeat Biopsy Study Groups

As stated previously, this investigation is similar in design to the other repeat biopsy studies it seeks to validate. Like others, it is a prospective cohort of male patients referred for repeat prostate biopsy because of persistent clinical suspicion for CaP. The patient population of U.S. veterans is similar to the populations found in the studies of Roehrborn et al. (1996) and Fowler et al.

(2000) (please see Table 1). A commonly used criterion for comparison of patient populations in repeat prostate biopsy studies is overall prevalence of CaP in the repeat group, or the proportion of false negatives from after the first prostate biopsy. This proportion may also reflect, in part, the similarities or differences in biopsy technique between study groups. The prevalence of CaP in this study was 22.6% (60/266), which is comparable to the prevalence of CaP in existing studies (please see Table 6). Roehrborn et al. (1996) found that the prevalence of CaP in their repeat biopsy patients was 23% (28/123)[11] and Fowler et al. (2000) found a prevalence of 27% (80/298)[20].

Table 6: False negative proportion at each repeat prostate biopsy

Number of Repeat Biopsies	False negative proportion
First repeat biopsy	23.7% (44/186)
Second repeat biopsy	21.3% (13/61)
Three of more repeat biopsies	15.8% (3/19)
Total	22.6% (60/266)

Other studies had false negative proportions between10% and 43%. The lower prevalences are due to differences in study design (i.e., repeat biopsies in all patients with negative initial biopsies[19] and the higher prevalences are due to different biopsy technique (transperineal saturation biopsies[21, 22]). Overall, the prevalence of CaP in the aggregate of patients with repeat prostate biopsies was 20.6% (639/3107). Both the overall prevalence and this investigation's finding are similar (X^2 = 0.616, df 1, p>0.5). Finally, the prevalence of 22.6% of CaP in this repeat biopsy study is very similar to the true false negative proportion of 23% found by Rabbani et al. (1998) in their investigation of false negative repeat sextant prostate biopsies in men already diagnosed with CaP.

A Kaplan & Meier (1958) survival curve[43] was generated to estimate the 1, 2 and 5 year incidence of CaP in the repeat prostate biopsy group (please see Appendix A: Figure 3). The median time of follow-up for these data was 4.8 years. The proportion being diagnosed with CaP at 1, 2 and 5 years are 0.1145, 0.1095 and 0.000, respectively (please see Table 7). The peak incidence of CaP diagnosis was at 4 years (0.2326).

Table 7: Life table for CaP diagnosis over time

Time (year)	Number exposed to risk	Proportion diagnosed	SE of surviving
0	227	0.1145	0.0211
1	137	0.1095	0.0302
2	80	0.1125	0.0387
3	44	0.1136	0.0479
4	22	0.2326	0.0674
5	5	0.0000	0.0674
6	1	0.0000	0.0674

Univariate Analyses Between Subjects with Positive and Negative Biopsies

Univariate analyses were carried out between the groups of patients who had positive and negative repeat prostate biopsies, as schematized in Figure 1. The results of the Pearson chi-square tests for categorical data and Mann-Whitney U tests for continuous data are summarized in Tables 8 and 9. Race, referral indication, prostate volume, PSA (as categorized into normal, elevated, and very elevated groups), and PSAD were found to be statistically significant to a two-tailed significance of $p < 0.05$ between those with and without CaP.

Specifically, those subjects diagnosed with CaP were more likely to be non-White (8.3% Black vs. 3.9% Black), have a referral indication of both elevated PSA and abnormal DRE (26.7% vs. 12.7%), have a smaller prostate gland (30.8 cm^3 vs. 39.35 cm^3), have elevated PSAs (greater than 4.0 ng/ml),

and have elevated PSA densities (0.215 vs. 0.155). Characteristics which did not distinguish between subjects with and without CaP included age, family history of CaP, prior vasectomy history, DRE, TRUS findings, number of PSAs, PSA velocity, PSA doubling time, presence of PIN, number of repeat prostate biopsy visits, first biopsy result, interbiopsy period, intertesting period, nor follow-up period.

Table 8: Significant differences between those with CaP and without CaP (negative)

Variable	P	Negative (N=206)	CaP (N=60)
Race	0.034		
White		93.6%	83.3%
Black		3.9%	8.3%
Hispanic		1.0%	6.7%
Referral Indication	0.044		
Abnl DRE		17.1%	8.3%
Elev. PSA		66.8%	63.3%
DRE & PSA		12.7%	26.7%
Initial Volume (median, cm^3)	0.001	39.35	30.80
Initial PSA - categorized	0.020		
0-3.99 ng/ml		17.5%	3.3%
4-9.99 ng/ml		65.0%	73.3%
>= 10 ng/ml		17.5%	23.3%
Initial PSAD (median, $ng/ml/cm^3$)	0.000	0.155	0.215

Table 9: Non-significant differences between those with CaP and without CaP (negative)

Variable	P	Negative (N=206)	CaP (N=60)
Age at entrance (median, years)	0.376	65.2	66.1
Family History of CaP	0.162		
No		82.3%	75.0%
Yes		17.7%	25.0%
Prior Vasectomy	0.755		
No		67.8%	65.0%
Yes		32.2%	35.0%
Initial DRE	0.247		
Not suspicious		65.0%	56.7%
Susp., Likely		35.0%	43.3%
Initial TRUS	0.075		
Not hypoechoic		57.3%	44.1%
Hypoechoic		42.7%	55.9%
PSA velocity (median, ng/ml/yr)	0.095	0.529	0.946
PSA doubling time (median, years)	0.074	6.50	5.52
Number of Biopsies	0.819		
Two		68.9%	73.3%
Three		23.3%	21.7%
Four		6.3%	3.3%
Five		1.5%	1.7%
Presence of PIN (ever)	0.211		
No		82.3%	75.0%
Yes		17.7%	25.0%
First Bx Result	0.728		
Normal		66.3%	66.1%
Inflammation		7.1%	5.4%
BPH		3.7%	5.4%
Atypia		4.8%	8.9%
Number of PSAs	0.301		
Two		34.5%	45.0%
Three		54.9%	50.0%
Four		9.3%	3.3%
Five		1.5%	1.7%
Interbiopsy period (median, months)	0.704	16.7	15.6
Intertesting period (median, months)	0.731	12.9	10.9
Follow-up period (median, months)	0.281	17.4	15.4

Multiple Logistic Regression Model

The variables found to be significant to a two-tailed significance of

P<0.100 in the univariate analyses were included in a multiple logistic regression model. Age of the patient, family history of CaP, DRE, and presence of PIN were also included into the model because of their clinical importance. The following variables were considered for incorporation into a multiple logistic model using a backwards stepwise procedure (maximum likelihood ratios): age, race, referral indication, TRUS findings, volume, PSA, PSAD, PSA velocity, PSA doubling time, DRE, presence of PIN, family history of CaP. Before the model was developed, Spearman's correlation coefficients were performed on continuous or ordinal data that met the assumptions of a correlation model to uncover possible multicollinearities between variables. The results are located in Table 10 in order of strength of correlation.

Table 10: Significant correlation coefficients between variables in the logistic model

Variables	N	Spearman's Correlation Coefficient	P
Volume X PSA	266	0.322	0.000
Age X Volume	266	0.263	0.000
Age X PSA	266	0.227	0.000

An interaction term was created for the significant correlation between volume and PSA and included in the model. However, this interaction term did not appear in the final model. The model also did not tolerate the addition of the variables of race and referral indication, as indicated by the large standard errors of their coefficients, due to the low numbers of subjects in many of the categories of these two variables. These variables were therefore excluded from the model. The final model is tabulated in Table 11.

The multiple logistic regression model had a chi-square statistic of 26.58 (df 4, p=0.000) and an overall correct classification of 76.2% (n=230). The

Hosmer and Lemeshow goodness-of-fit test chi-square was 6.17 (df 8, p=0.62) which shows the regression equation has an adequate fit to the data, i.e., the observed and expected results are similar[37]. PSA, as categorized into low (0-3.99 ng/ml), elevated (4-9.99 ng/ml) and highly elevated (10 ng/ml) was the most predictive variable with an odds ratio of 13.9 (95% CI 2.5-76.5) for PSA values between 4 and 9.99 ng/ml. Prostate volume was protective with an odds ratio of 0.96 (95% CI 0.94-0.98). Abnormal DRE findings were also associated with a positive final repeat biopsy, but this finding was not statistically significant (p=0.603).

Table 11: Logistic model for predicting positive or negative repeat prostate biopsies

Variable	B	S.E.	Sig.	Exp (B)	95% CI	for Exp (B)
PSA - categorized			0.0099			
PSA 0-3.99 ng/ml	2.165	0.7795	0.0055	8.7188	1.8921	40.1767
PSA 4-9.99 ng/ml	2.631	0.8703	0.0025	13.8984	2.5247	76.5115
Volume	-0.037	0.0110	0.0008	0.9637	0.9431	0.9848
DRE (0)	0.633	0.3369	0.0603	1.8833	0.9730	3.6453
Constant	-2.031	0.8424	0.0159			

In influence regression diagnostic testing, there was one outlier at a Cook's distance of 1.18 (please see Figure 4), which is worrisome for assuming a correct model fit[33]. In examining the dataset, there were no apparent errors in data entry or calculation, and in a brief post hoc analysis, two more interaction terms (age*volume and age*PSA) were added with the result of an overall poorer fit of the model to the data. The Studentized residual plot (please see Figure 5) showed an unusual distribution of residuals with some hovering above the cut-off of 2, which is also a cause for concern, but could not be corrected[33].

Cox Proportional Hazards Regression Model

In addition to a multiple logistic regression model, a Cox proportional hazards model was fitted to the dataset. Statistically significant variables (two tailed, $p < 0.05$) in univariate and multivariate analyses in prior repeat prostate biopsy studies were incorporated into a Cox proportional hazards model using this data set. As stated in the Subjects, Apparatus, Procedure section, time was defined as the period between first and last biopsy, diagnosis of CaP was the terminating event and those subjects with negative last biopsies were censored. The variables under consideration for incorporation included: DRE result, number of repeat biopsy visits, PSA, PSAD, PSA velocity, volume and age.

Before entrance into the model, the variables that met the assumptions of correlation models were tested for possible multicollinearities in the model using pairwise Spearman's rank correlation coefficients. Pairwise correlations with time were also assessed and tabulated in Table 12. Variables with interactions of moderate to large effect were incorporated into the model using interaction terms in order to develop a thoughtful model without use of a correlation matrix.

Table 12: Significant correlation coefficients between covariates in the proportional hazards model

Variables	N	Spearman's Correlation Coefficient	P
# of Bxs X Time	266	0.400	0.000
Time X PSAD	266	-0.342	0.000
Volume X PSA	266	0.322	0.000
Time X PSA	266	-0.239	0.000
# of Bxs X Age	266	-0.206	0.001
Time X Volume	266	0.188	0.002

The independent variables were also subjected to Kaplan-Meier analyses

to test the proportional hazards assumption: the hazard ratio continues to be constant for any fixed set of the independent variables over time. The survival curves of the independent variables were plotted and those that violated this assumption by crossing were included into the model after controlling for time-dependence. Thusly, time-dependent covariates were created for the variables DRE result (T_DRE), PSA velocity (T_PSAV) and number of sets of biopsies (T_BXSETS).

As in the multiple logistic model, an interaction term was created for the pairwise correlation between volume and PSA, but did not appear in the final model. The time-dependent variable PSA velocity x time also did not appear in the model, and to improve the proportional hazards model's ability to correctly describe the data, this variable was excluded and the subject number in the final model was 259 (other variables had missing data). The Cox proportional hazards regression model had an overall chi-square statistic of 43.09 (df 3, p=0.000) with a -2 log likelihood of 535.97. The final model is tabulated in Table 13.

Table 13: Proportional hazards regression model for predicting positive or negative repeat biopsies

Variable	B	S.E.	Sig.	Exp (B)	95% CI	for Exp (B)
PSAD	2.1803	0.4950	0.0000	8.8486	3.3537	23.3468
Volume	-0.0197	0.0083	0.0182	0.9805	0.9646	0.9966
T_BXSETS	-0.0111	0.0070	0.1159	0.9890	0.9754	1.0027

Overall, the model had a sufficient fit to the data with a very significant model chi-square and significance. PSA density was, by far, the most predictive variable with an odds ratio of 8.85 (95% CI 3.35-23.35). Volume had a similar result as the multiple logistic regression model with an odds ratio of 0.98 (95% CI

0.96-0.997). Finally, the time-dependent variable correcting the number of biopsy sets was also in the model, but was not statistically significant. Regression diagnostics showed that the model was adequate in describing the data with the greatest |DfBetai,j| being considerably less than 1[50] (please see Figure 6). The largest DfBetai,j was a single point at 0.416 with the vast majority of points hovering near zero.

Area Under the ROC Curves

The performance of the variables that were statistically significant in the multiple logistic model or Cox proportional hazards model was determined using area under the receiver operating characteristics curve (ROC) analyses[38]. Please see Figures 7, 8 and 9 for the ROC plots of PSAD with volume, PSA categorized with volume, and PSAD. Table 14 summarizes the results of the ROC analyses for statistically significant covariates and selected interactions of the two regression models. The highest performing classifier was PSA density interacting with volume: ROC of 0.693 (S.E. = 0.037, p=0.000). The highest single performing classifier was PSAD with an ROC of 0.681 (S.E.= 0.037, p=0.000). Other statistically significant ROCs include PSA-categorized interacting with volume, and volume.

Table 14: Area Under the Curve (ROC) analyses for model covariates

Variable	Area	Std. Error	Sig.	95% CI Lower Interval	Upper Interval
PSAD with volume	0.693	0.040	0.000	0.614	0.772
PSA - categorized with volume	0.683	0.040	0.000	0.604	0.762
PSAD	0.681	0.037	0.000	0.609	0.753
Volume	0.640	0.038	0.001	0.566	0.714
PSA - categorized	0.583	0.039	0.052	0.505	0.660

a. The test result variables have at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.
b. Under the nonparametric assumption
c. Null hypothesis: true area = 0.5

Operating points were chosen by finding the point on the ROC curves closest to the northwest corner - or the coordinate where true positives are 100% and false positives are 0%[39]. The operating points, sensitivities and specificities are tabulated for the covariates and selected interactions in Table 15. The relationship between PSAD and volume was defined by calculating PSAD depending on the operating point of volume, which was 40.45 cm^3. Because volume was negatively associated with CaP in the models, if the volume was greater than 40.45 cm^3, then PSAD was calculated to swap the polarity of the classifier, effectively rotating the ROC curve around 40.45 cm^3. The relationship between categorized PSA and volume was handled in the same manner.

Table 15: Operating points ("cut-offs") of the model covariates

Variable	Operating Point	Sensitivity	Specificity
PSAD with volume ($ng/ml/cm^3$)	0.135	0.750	0.646
PSA - categorized with volume	7.050	0.750	0.631
PSAD ($ng/ml/cm^3$)	0.135	0.850	0.422
Volume (cm^3)	40.45	0.817	0.481
PSA with volume (continuous, ng/ml)	4.55	0.936	0.377

Prostate volume was negatively associated with CaP in both the multiple logistic model and the proportional hazards model, but the 95% confidence intervals for e^B skirted unity. However, if prostate volume is categorized by the cut-off value of 40.45 cm^3, a simple logistic regression shows the odds ratio to be 0.2426 (95% CI 0.1194-0.4929), a finding that is more consistent with the clinical significance of prostate volume (as demonstrated by the improvement in area under the ROC curve for PSAD and PSA).

Improving the Negative Predictive Value of PSA Indices

The sensitivities, specificities and negative predictive values (it is more important to prevent false negatives than false positives) of the covariates using their operating points were calculated (please see Table 16). Interestingly, the highest sensitivity and negative predictive value was obtained by using the multiple logistic model covariates: 100% and 100%, respectively, despite the lower area under the ROC curve of PSA and prostate volume compared with PSAD and prostate volume. However, only 4 subjects of the 206 subjects with repeat negative last biopsies met the classification criteria with an initial PSA < 4.0 ng/ml, a prostate > 40.45 cm^3 and a negative DRE. The reason for this low number is simple: the vast majority of subjects were referred for a repeat prostate biopsy if their initial PSA > 4.0 ng/ml and/or they had an abnormal DRE.

Table 16: Sensitivity, specificity and negative predictive values of the model covariates

Model	Covariates	Sensitivity	Specificity	- PV	False Negatives
Multiple Logistic	PSA, volume, DRE	100% (20/20)	15.4% (4/26)	100% (4/4)	0% (0/4)
Proportional Hazards	PSAD, volume	94.8% (73/77)	54.6% (53/97)	93% (53/57)	5.2% (4/57)

The proportional hazards model covariates, using the cut-offs of PSAD of 0.135 ng/ml/cm^3 and prostate volume of 40.45 cm^3, had a sensitivity of 94.8% and a negative predictive value of 93.0%. Fifty-seven subjects were classified with an initial PSAD <0.135 ng/ml/cm^3 and an initial prostate volume > 40.45 cm^3, 4 were false negatives.

Discussion

Design Limitations

The design of this investigation is prospective, but because the population is a clinically-based cohort, there is a significant selection bias towards those individuals the clinicians (both referring and specialist physician) felt it was important to follow. Less than 30% of the initially negative cohort was evaluated for repeat biopsy, creating a clinically selected sample (please see Figure 1). An unfortunate consequence of this clinically selected sample is that the true relative risk cannot be assessed. The CaP prevalence of the entire referred population is unknown: veterans who were not referred for repeat biopsy after an initially negative biopsy were not followed and consequently, their prevalence of CaP was not evaluated. There have been studies which have re-biopsied all individuals negative after the first biopsy, but as a consequence, the sample number was small (n= 54)[28] or the length of time between biopsies was inadequate to accurately determine a serum PSA (t=6 weeks)[19] given the recent prostate gland manipulation[45].

In addition to the clinical selection bias, the prospective cohort is not a true, active cohort because of the lack of active follow-up: 101 of the subjects (45.5%) found negative on the first repeat biopsy were not biopsied for a second repeat biopsy, and 38 of the subjects (47.5%) found negative on second, third and fourth biopsies were not biopsied again. Because of the highly selected nature of the patient population in this study, applications of the results of this

study must be limited to populations similar to the one in this study. Fortunately, these limitations of a patient population are commonly seen in the clinical setting encountered by a community urologist.

Subjects were selected for follow-up based on an abnormal DRE and/or elevated PSA, or presence of PIN, so variables dependent on those selection criteria are biased: DRE findings, PSA, PSAD, presence of PIN. Table 4 shows the degree to which these characteristics of subjects in the repeat group differ from the non-repeat group, a difference which may or may not have clinical significance. The median PSA and median PSAD are statistically significantly higher in the repeat group, and the median PSAD difference has some clinical relevance: a PSAD of 0.15 ng/ml/cm^3 has been determined to be associated with the presence of CaP[46].

Importantly, although the repeat prostate biopsy group was a clinically selected population, the degree to which it differed from the non-repeat prostate group is not catastrophic since the differences from a clinical perspective are not great (please see Table 4). Using the median, the difference in age was two years, the difference in prostate volume was 2.3 cm^3, the difference in PSA was 0.5 ng/ml, and the difference in PSA density was 0.02 ng/ml/cm^3. The difference in abnormal digital rectal exam is concerning, but the bias is towards a less clinically suspicious population (51.2% non-repeat, 37.1% repeat group). Finally, there is a difference between those with PIN (8% non-repeat, 12.1% repeat), which is a clinically important difference and a bias towards a more clinically suspicious group.

Rietbergen et al. (1998)[16] also compared their non-repeat biopsy group

and their repeat group, however, the only published statistic was the median volume of these groups, 37.5 cm^3 (interquartile range 6.4-224.6) and 48.2 cm^3 (interquartile range 10.0-175.5), respectively. The non-repeat and repeat groups of this investigation differed in prostate volume (33.5 cm^3 vs. 35.8 cm^3), but the difference was not as dramatic as shown by Rietbergen et al. (1998). The other comparisons made by Rietbergen et al. (1998) included PSA, prostate volume and PSA velocity, but these comparisons were made between initially positive subjects versus repeat positive subjects. These comparisons may or may not be valid because their repeat subjects were preferentially selected for higher PSA values (40% or 196/486 were not repeatedly biopsied because their PSA value was below 4 ng/ml)[16].

Comparing Results

The inflation of values of the PSA indices and the greater likelihood of abnormal DRE findings of the subjects relative to the non-repeat group may have made it more difficult to find statistically significant differences in these parameters. The entire repeat population has a lower likelihood of decreased PSA indices and normal DRE findings; therefore, there was an expectation that finding a difference in the variance in the population based on these parameters would be more difficult given this regression towards the null.

Despite the expectation of a regression towards the null, the findings of this investigation indicate a strong predictive value of PSA density (p=0.0000) interacting with prostate volume (p=0.0182). In addition to these parameters,

another covariate included in the Cox proportional hazards model is the time-dependent covariate number of prostate biopsy sets (p=0.1159). As stated earlier, the area under the ROC curve of PSAD interacting with volume was 0.693 (p=0.000, with swap in polarity) with an operating point of PSAD = 0.135 and prostate volume of 40.45 cm^3.

Clinically, the meaning of a swap in polarity is that if a patient has a PSA density of greater than 0.135 ng/ml/cm^3 and a prostate gland smaller than 40.45 cm^3, then he is at greater risk for a subsequent positive repeat positive biopsy according to the Cox proportional hazards model. The operating point of categorized PSA and volume has less clinical meaning, but returning to the multiple logistic model, PSA values from 4-9.99 ng/ml, and $\geq$ 10 ng/ml, with a prostate gland < 40.45 cm^3, place a patient at greater risk for a positive repeat prostate biopsy. The operating points of prostate volume and PSA are more straightforward to interpret clinically. A prostate volume > 40.45 cm^3 is protective, and a PSA value > 4.55 ng/ml places a patient at greater risk for a positive biopsy.

In comparison to the sensitivities, specificities and negative predictive values found in Table 16, Ukimura et al. (1997) also published these values for PSA indices. They define an entity called volume-referenced PSA[14] which considers different PSA values as abnormal for different sizes of the prostate, e.g., PSA > 2.6 ng/ml for 25 cm^3 < volume $\leq$ 35 cm^3, PSA >3.4 ng/ml for 35 cm^3 < volume $\leq$ 45 cm^3. This entity is the somewhat similar to the multiple logistic model, although the cut-off for PSA was higher in this investigation (4.0 ng/ml). The sensitivity, specificity and negative predictive values for volume-referenced PSA were 85%, 21% and 87%, respectively. They do not define a "volume-referenced PSAD," but their sensitivity, specificity and negative predictive value

for PSAD was 79%, 48% and 95%. Both volume-referenced PSA and PSAD had high false negative proportions, 15% and 21%, which is higher than the 0% and 5.2% found in this investigation.

However, PSA velocity was not found to be significant in this investigation as it had been in the studies of Perachino et al. (1997)[13], Ukimura et al. (1997)[14] and Borboroglu et al. (2000)[18]. This may be a consequence of the regression towards the null (almost all subjects were preferentially selected for rising PSAs and finding a difference between those finally positive for CaP versus those negative may be difficult). Alternatively, many of the calculations of PSA velocity in this study were calculated with short interbiopsy periods, and as PSA velocity is inversely dependent on interbiopsy time, there may have been miscalculation of the true regression of the rise in PSA. An alternative procedure for obtaining the slope of PSA elevation has been proposed by Keetch et al. (1996)[9], which is not as prone to distortion by time, but has been criticized for being difficult to calculate for a clinical urologist.

The findings of this investigation confirm the findings of some previous repeat biopsy studies, although not perfectly. In the multiple logistic model, PSA, volume, and DRE findings were covariates of the regression model. PSA was found to have an area under the ROC curve of 0.583 ($p=0.052$), a finding consistent with the poor performance of the area under the ROC curve for PSA found by Roehrborn et al. (1996) (0.63, $p=0.02$)[11]. Prostate volume was found to be statistically significant as a protective parameter, as it has in several other repeat biopsy studies[16, 17, 21]. Volume may be negatively associated with a final

diagnosis of CaP because of the difficulty in finding small cancers in large prostates (Bayesian theory), or because of the interaction with benign prostate hyperplasia and prostatitis with elevated PSAs - patients selected for elevated PSAs with these two benign entities would have a lower prevalence of CaP.

DRE was not found to be statistically significant (OR= 1.88, 95% CI 0.97-3.65) in the multiple logistic model although it was in the hierarchical multiple logistic regression modeled by Keetch et al., 1994 (OR= 3.5)[9]. The Cox proportional hazards model validated Fowler et al.'s (2000)[20] finding that PSAD was a predictive covariate: they found the area under the ROC for PSAD to be 61.8%, which is considerably less than the area found in this investigation (68.1% or 69.3% with prostate volume).

The finding that the number of prostate biopsy sets was associated with the prostate biopsy result (OR=0.99, 95% CI 0.98-1.00) was also found by Keetch et al., (1994)[9], although the odds ratio was different (OR=0.2). This difference may be attributable to the relationship between time and the number of prostate biopsy sets. In Figure 10, the survival curves of the subjects found with CaP at different biopsy numbers is shown. It is apparent that subjects with 2 prostate biopsies (or one repeat set of biopsy) have the steepest survival curve, but since the hazard ratios are not constant, the survival curves of subjects with 3, 4 and 5 prostate biopsy sets cross. In this investigation, the peak incidence of CaP was at four years, median time of follow-up was 4.8 years. These observations are likely secondary to the passive nature of the follow-up of this cohort - subjects were free to drop out or were censored by the primary care

provider or urologist.

Model Comparisons

The multivariate model exclusively used by other repeat prostate biopsy studies was the multiple logistic regression model. This study also used a proportional hazards model to confirm the findings of other repeat prostate biopsy studies. Comparing the assumptions inherent to using a multiple logistic regression model or a proportional hazards regression model, a logistic regression model assumes that there is a random, independently selected population (or representative or the larger population) and that there are no interactions between covariates. A proportional hazards regression model has the same assumptions with the additional, crucial, assumption that the hazard ratios of the covariates remain constant over time. The population of this investigation more or less conforms to these assumptions; where there were possible severe multicollinearities, as suggested by high pairwise correlation coefficients, or with variable hazard ratios, the models were changed in order to comply.

However, this population has characteristics that may make it unsuitable for analysis by a logistic regression model or a proportional hazards model. As alluded to earlier, because subjects are involved for variable lengths of time, there may be a misclassification of shorter term subjects who are diagnosed as negative at the time the analyses were performed, but may have become positive if they had been followed longer. The models were somewhat different, which suggests that time was a confounder. Also, because the proportional hazards

regression model is a survival analysis, there is an assumption that censoring is random. Unfortunately, because data were not gathered as to why subjects were censored, this assumption of the proportional hazards model cannot be addressed.

In comparing the multiple logistic and proportional hazards models as a function of their fitness and adequacy in describing the data, the multiple logistic regression model was unable to adequately describe the data without notable violations of Cook's statistic or the standard error of the Studentized residual. The Cox proportional hazards model with time-dependent variables was able to fit the data adequately without violations as diagnosed by the DfBeta statistic. The proportional hazards model was also able to describe the independent variance of prostate volume along with the variance attributable to PSAD, a finding that improved the area under the ROC curve of the covariates of the proportional hazards model over the covariates of the multiple logistic model.

Data Limitations

The greatest limitations of this investigation relative to other studies of repeat prostate biopsies include the absence of %-free-to-total PSA measurements, the omission of transition zone volume estimations and the lack of systematic transition zone prostate biopsies. In other repeat prostate biopsy studies, %-free-to-total PSA was found to be the most important variable associated positive repeat biopsy[19, 20]. Fowler et al., (2000) report an area under the ROC curve for %-free-to-total PSA as 74.5%[20], which bests the covariates of

these analyses (please see Table 13). However, the association of this parameter as a risk factor for positive repeat biopsy is not consistently significant. For example, Borboroglu et al. (2000)[18] did not find %-free-to-total PSA to be significant in their univariate analyses. It would have been useful to replicate either of these findings. The lack of transition zone-related parameters is also notable, as there might be an underestimation of the true prevalence of CaP. Indeed, Djavan and colleagues (1999)[48] found that sampling the transition zone systematically yielded greater prevalence of CaP. The counterargument is that the omission of transition zone studies may be more clinically relevant with regards to "significant" cancers[49].

Summary and Conclusion

This investigation may not be the definitive repeat prostate biopsy study, but it is a novel approach to this question and contributes more information. PSA density, prostate volume, the number of repeat biopsies and DRE were confirmed as risk factors previously found to be associated with a positive repeat prostate biopsy. PSA velocity was not found to be a statistically significant predictor in this cohort. Using a Cox proportional hazards model in this study design was a valuable tool in the assessment of risk for a positive final repeat prostate biopsy.

These findings are clinically relevant to a practicing urologist. As stated previously in the introduction, the question of whom to select for a repeat prostate biopsy has not been consistently answered in the medical literature. This investigation synthesizes information and contributes data that confirms some groups' findings. Specifically, if a man has an initially negative prostate biopsy but is clinically indicated to return for a repeat biopsy (due to elevated PSA indices) he is at decreased risk for a final prostate biopsy if he has an initial PSA density < 0.135 ng/ml/cm^3 with a prostate gland > 40 cm^3. If his PSA is also < 4.0 ng/ml, then he is at even less risk for a positive final prostate biopsy. The initial digital rectal examination, an important tool in the clinical evaluation of a man suspected of having prostate cancer and also a key component of prostate cancer staging[47], was also confirmed as a risk factor for a positive final repeat prostate biopsy.

Future directions in the evaluation of the risk for a positive repeat prostate

biopsy should address the selection biases in a clinical sample, should actively follow cohorts over time, should incorporate newer technologies shown to be associated with CaP (such as %-free-to-total PSA) and most importantly, should address the significance of the "insignificant" cancer more likely to be diagnosed with repeated biopsies.

References

1. Howe, H. L., Wingo, P. A., Thun, M. J., et al.: Annual report to the nation on the status of cancer (1973 through 1998), Featuring cancers with recent increasing trends. JNCI **93**(11): 824-842, 2001.

2. Dugan, J. A., Bostwick, D. G., Myers, R. P., et al.: The definition and preoperative prediction of clinically insignificant prostate cancer. JAMA **275**(4): 288-294, 1996.

3. Gleason, D. F. and Mellinger, G. T.: Prediction of prognosis for prostatic adenocarcinoma by combined histologic grading and clinical staging. J Urol **111**: 58, 1974.

4. Stamey, T. A., McNeal, J. E., Yemoto, C. M., et al.: Biological determinants of cancer progression in men with prostate cancer. JAMA **281**(15): 1395-1400, 1999.

5. Stamey, T. A.: Making the most out of six systematic sextant biopsies. Urology **45**: 2-12, 1995.

6. Shibata, A. and Whittemore, A. S.: Re: Prostate cancer incidence and mortality in the United States and the United Kingdom. JNCI **93**(14): 1109-1110, 2001.

7. Hodge, K. K., McNeil, J. E. and Terris, M. K.: Random systematic versus directed ultrasound guided transrectal core biopsies of the prostate. J Urol **142**: 71, 1989.

8. Rabbani, F., Stroumbakis, N., Kava, B. R., et al.: Incidence and clinical significance of false-negative sextant prostate biopsies. J Urol **159**: 1247-1250, 1998.

9. Keetch, D. W., Catalona, W. J. and Smith, D. S.: Serial prostatic biopsies in men with persistently elevated serum prostate antigen values. J Urol **151**: 1571-1574, 1994.

10. Ellis, W. J. and Brawer, M. K.: Repeat prostate needle biopsy: who needs it? J Urol **153**: 1496-1498, 1994.

11. Roehrborn, C. G., Pickens, G. J. and Sanders, J. S.: Diagnostic yield of repeated transrectal ultrasound0guided biopsies stratified by specific histopathological diagnoses and prostate-specific antigen. Urology **47**: 347, 1996.

12. Fleshner, N., O'Sullivan, M. and Fair, W.: Prevalence and predictors of a positive repeat transrectal ultrasound-guided needle biopsy of the prostate. J Urol **158**: 505-509, 1997.

13. Perachino, M., Ciolo, L. d., Barbetti, V., et al.: Results of rebiopsy for suspected prostate cancer in symptomatic men with elevated PSA levels. European Urology **32**(q): 155-9, 1997.

14. Ukimura, O., Durrani, O. and Babaian, R. J.: Role of PSA and its indices in determining the need for repeat prostate biopsies. Urology **50**(1): 66-72, 1997.

15. Letran, J. L., Blase, A. B., Loberiza, F. R., et al.: Repeat ultrasound guided prostate needle biopsy: Use of free-to-total prostate-specific antigen ratio in predicting prostatic cancer. J Urol **160**: 426-429, 1998.

16. Rietbergen, J. B., Kruger, A. E. B., Hoedemaeker, R. F., et al.: Repeat screening for prostate cancer after 1 year followup in 984 biopsied men: Clinical and pathological features of detected cancers. J Urol **160**: 2121-2125, 1998.

17. Epstein, J. I., Walsh, P. C., Akingba, G., et al.: The significance of prior benign needle biopsies in men subsequently diagnosed with prostate cancer. J Urol **162**: 1649-1652, 1999.

18. Borboroglu, P. G., Comer, S. W., Riffenburgh, R. H., et al.: Extensive repeat transrectal ultrasound guided prostate biopsy in patients with previous benign sextant biopsies. J Urol **163**: 158-162, 2000.

19. Djavan, B., Zlotta, A., Remzi, M., et al.: Optimal predictors of prostate cancer on repeat biopsy: a prospective study of 1,051 men. J Urol **163**: 1144-1149, 2000.

20. Fowler, J. E., Bigler, S. A., Miles, D., et al.: Predictors of first repeat biopsy cancer detection with suspected local stage prostate cancer. J Urol **163**: 813-811, 2000.

21. Igel, T. C., Knight, M. K., Young, P. R., et al.: Systematic transperineal ultrasound guided template biopsy of the prostate in patients at high risk. J Urol **165**: 1575-1579, 2001.

22. Stewart, C. S., Liebovich, B. C., Weaver, A. L., et al.: Prostate cancer diagnosis using a saturation needle biopsy technique after previous negative sextant biopsies. J Urol **166**: 86-92, 2001.

23. Carter, H. B., Pearson, J. D., Waclawiw, Z., et al.: Prostate-specific antigen variability in men without prostate cancer: effect of sampling interval on prostate-specific antigen velocity. Urology **45**(4): 591-596, 1995.

24. Babaian, R. J., Toi, A., Kamoi, K., et al.: A comparative analysis of sextant and an extended 11-core multisite directed biopsy strategy. J Urol **163**: 152-157, 2000.

25. Stricker, H. J., Ruddock, L. J., Wan, J., et al.: Detection of non-palpable prostate cancer. A mathematical and laboratory model. British J Urol **71**: 43, 1993.

26. Eskew, L. A., Bare, R. L. and McCollough, D. L.: Systematic 5 region prostate biopsy is superior to sextant method for diagnosing carcinoma of the prostate. J Urol **157**: 199, 1997.

27. Levine, M. A., Ittman, M. and Melamed, J.: Two consecutive sets of transrectal ultrasound guided sextant biopsies of the prostate for the detection of prostate cancer. J Urol **159**: 471, 1998.

28. Carter, H. B., Pearson, J. D., Metter, J., et al.: Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. JAMA **267**(16): 2215-2220, 1992.

29. McNeal, J. E.: Normal histology of the prostate. Am J of Surg Pathol **12**: 619, 1988.

30. Stamey, T., Yang, N. and Hay, A. R.: PSA as a serum marker for adenocarcinoma of the prostate. NEJM **317**: 909, 1987.

31. Terris, M. K. and Stamey, T. A.: Determination of prostate volume by transrectal ultrasound. J Urol **145**: 984, 1991.

32. Schmid, H. P., McNeal, J. E. and Stamey, T. A.: Observations on the doubling time of prostate cancer. Cancer **71**(6): 2031-2040, 1993.

33. Glantz, S. A. and Slinker, B. K.: Primer of applied regression & analysis of variance. New York: McGraw-Hill, Inc., 2001.

34. Cox, D. R.: Regression models and life tables. J of the Royal Statistical Society **34**: 187-220, 1972.

35. Katcher, J., Kupelian, P.A., Zippe, C., Klein, E.A., et al.: Indications for excluding the seminal vesicles when treating clinically localized prostatic adenocarcinoma with radiotherapy alone. International Journal of Radiation Oncology, Biology, Physics. **37**(4):871-6, 1997.

36. Stamey, T. A., McNeal, J. E., Yemoto, C. M., et al.: Biological determinants of cancer progression in men with prostate cancer. JAMA **281**(15): 1395-1400, 1999.

37. Selvin, S.: Statistical analysis of epidemiological data. New York: Oxford University Press, 1991.

38. Hanley, J. A. and McNeil, B. J.: The meaning and use of the area under a receiver operating characteristics (ROC) curve. Radiology **143**: 29-36, 1982.

39. Swets, J. A., Dawes, R. M. and Monahan, J.: Psychologic science can improve diagnostic decisions. Psychologic science in the public interest **1**: 1-26, 2000.

40. Rose, C. C., Murphy, J. G. and Schwartz, J. S.: Performance of an index predicting the response of patients with acute bronchial asthma to intensive emergency department treatment. NEJM **310**: 573-577, 1984.

41. Kaye, K. W. and Lightner, D. J.: Prostate biopsy indices: toward efficient use of transrectal ultrasound. Urology **41**: 417, 1993.

42. McNeal, J. E., Villers, A. A., Redwine, E. A., et al.: Histologic differentiation, cancer volume, and pelvic node metastases in adenocarcinoma of the prostate. Cancer **66**: 1225, 1990b.

43. Johansson, J. E., Adami, H. O., Andersson, S. O., et al.: High 10-year survival rate in patients with early, untreated prostatic cancer. JAMA **267**: 2191, 1992.

44. Kaplan, E. L. and Meier, P.: Nonparametric estimation from incomplete observations. J Am Stats Assoc **53**: 457-485, 1958.

45. Yuan, J. J., Coplen, D. E. and Petros, J. A.: Effects of rectal examination, prostatic massage, ultrasonography, and needle biopsy on serum PSA levels. J Urol **147**: 810, 1992.

46. Benson, M. C., Whang, I. S., Olsson, C. A., et al.: The use of prostate specific antigen density to enhance the predictive value of intermediate levels of prostate specific antigen. J Urol **147**: 817-821, 1992.

47. Presti, J. C. (2000). Neoplasms of the prostate gland. Smith's General Urology. E. A. Tanagho and J. W. McAninch. New York, Lange Medical Books/ McGraw-Hill.

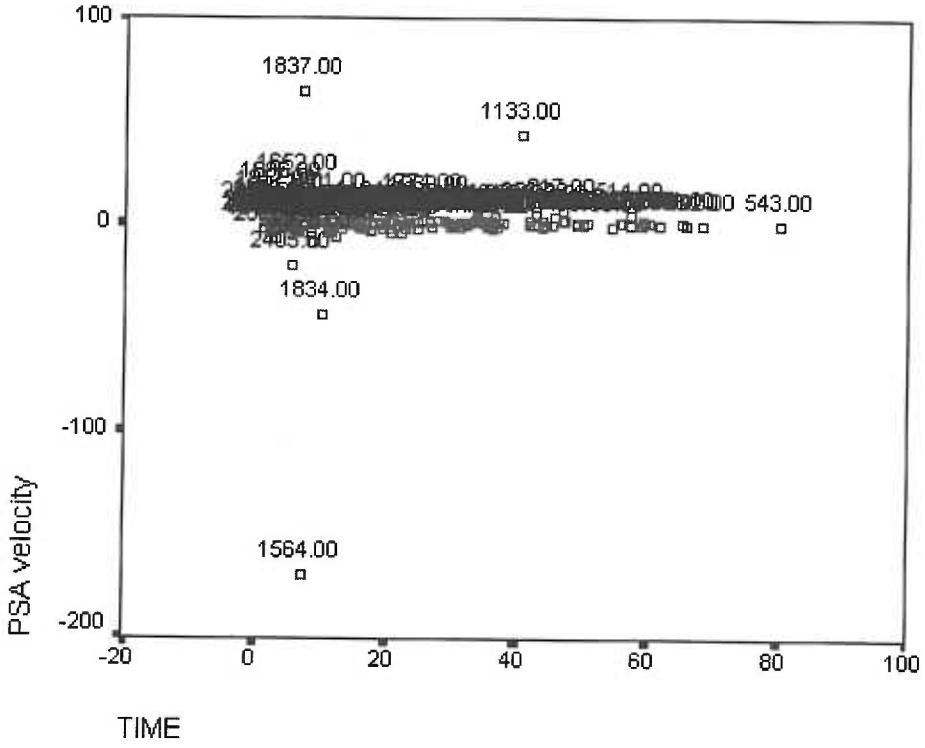
48. Djavan, B., Zlotta, A., Byttebier, G., et al.: Prostate specific antigen density of the transition zone for early detection of prostate cancer. J Urol **160**: 411-419, 1998.

49. Terris, M. K., Pham, T. Q., Issa, M. M., et al.: Routine transition zone and seminal vesicle biopsies in all patients undergoing transrectal ultrasound guided prostate biopsies are not indicated. J Urol **157**: 204, 1997.

50. Bollen, K.A. and Jackman, R.W. (1990). Regression diagnostics: an expository treatment of outliers and influential cases. Modern Methods of Data Analysis. J. Fox and J.S. Long. Newbury Park, CA. Sage Publications.

Appendix A: Figures 2-10

Figure 2: PSA Velocity Outliers (PSA velocity -ng/ml/year, Time -months)



Appendix A: Figures 2-10
(Continued)

Figure 3: Survival Curve for Diagnosis of CaP

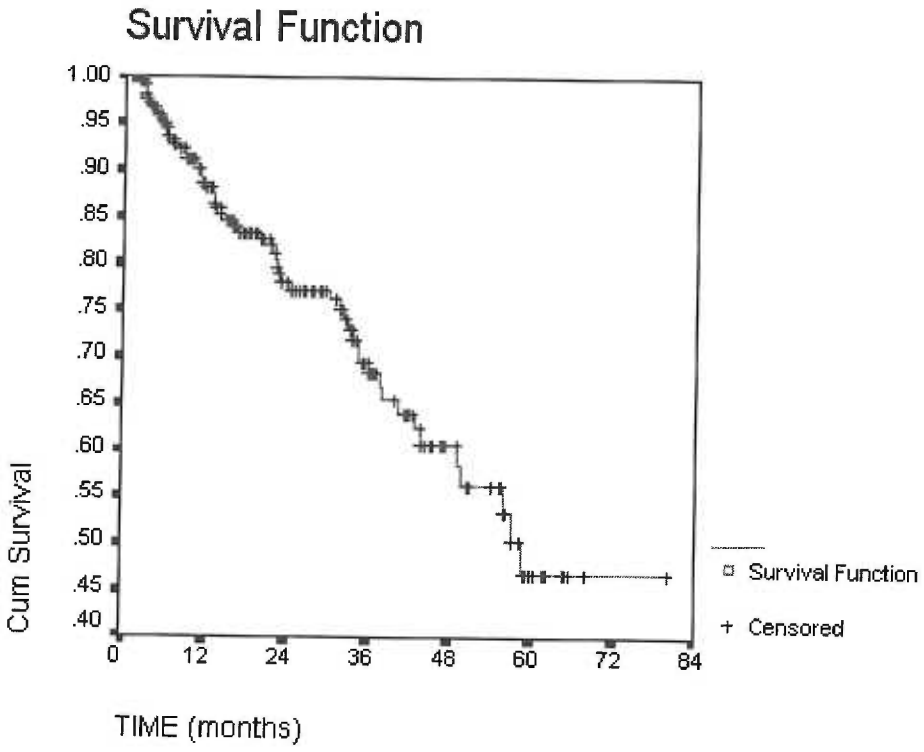
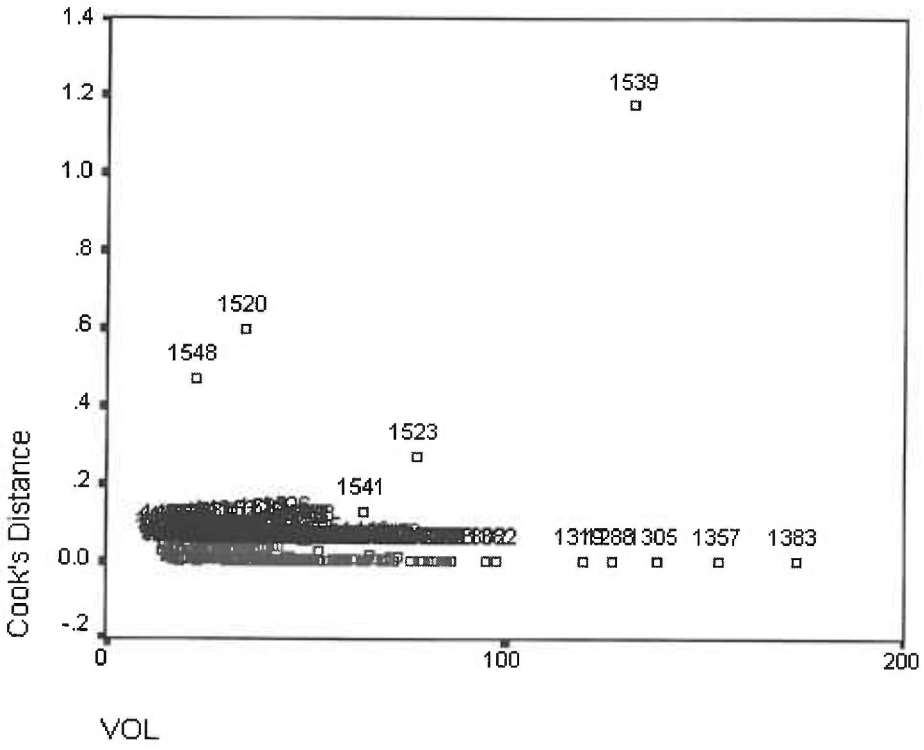
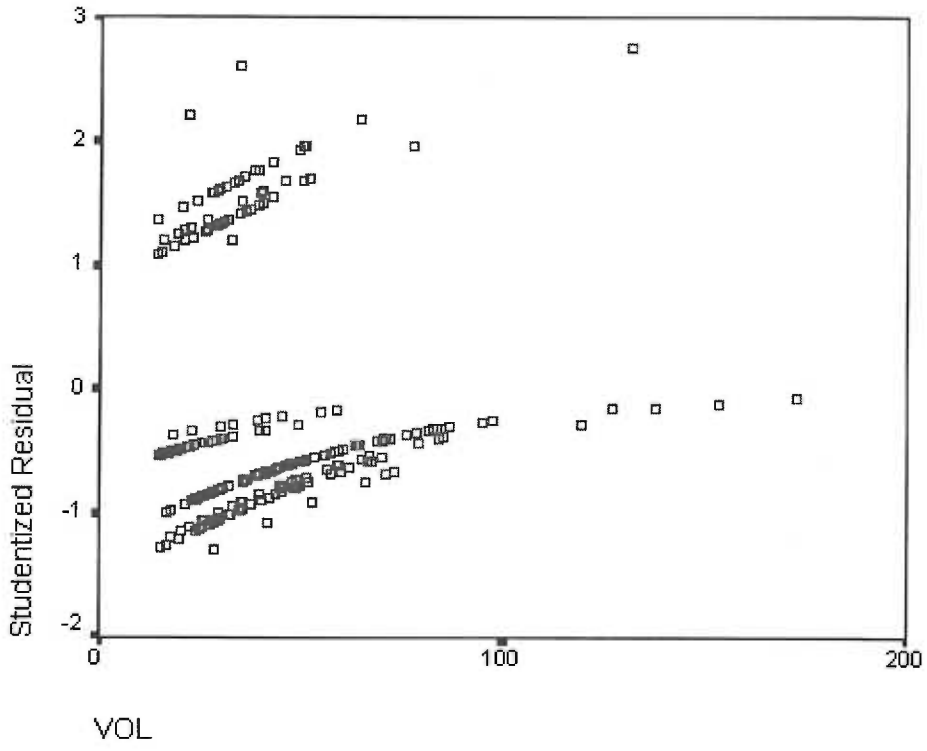


Figure 4: Cook's Distances for Multiple Logistic Regression Model



Appendix A: Figures 2-10
(Continued)

Figure 5: Studentized Residuals for Multiple Logistic Regression Model



Appendix A: Figures 2-10
(Continued)

Figure 6: $DfBeta_{i,j}$ Statistics for the Proportional Hazards Model

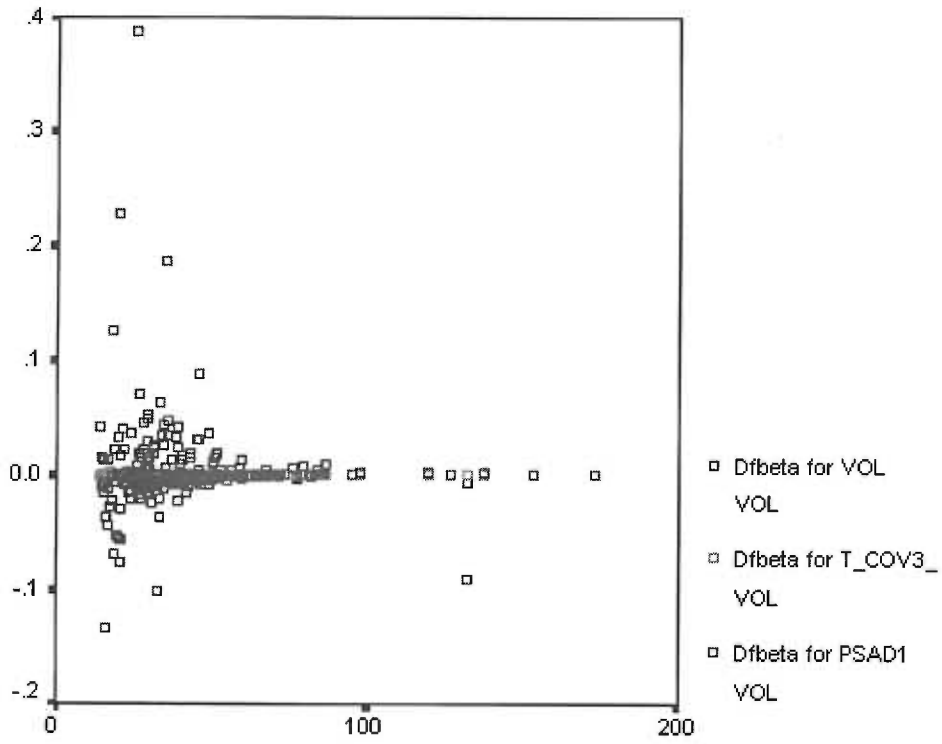


Figure 7: ROC Curve for PSA Density and volume

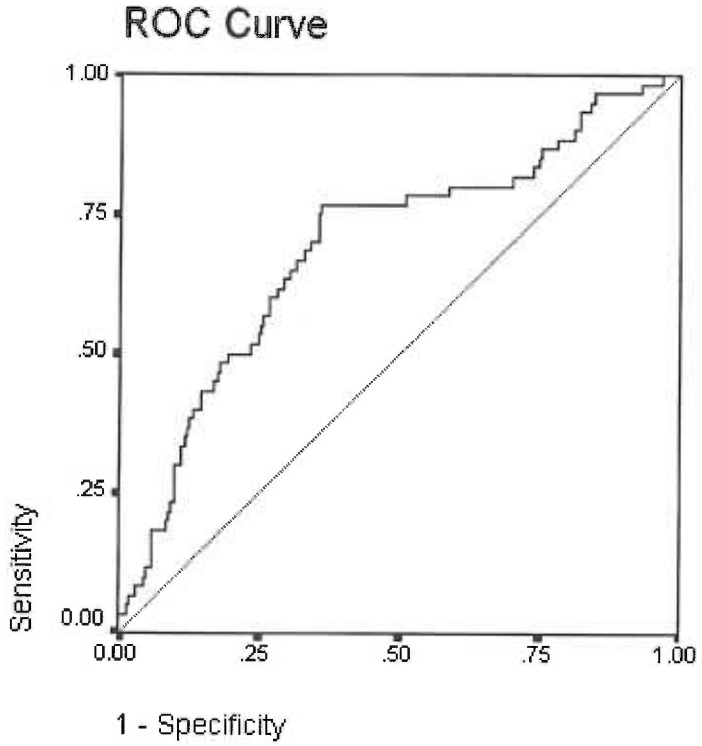
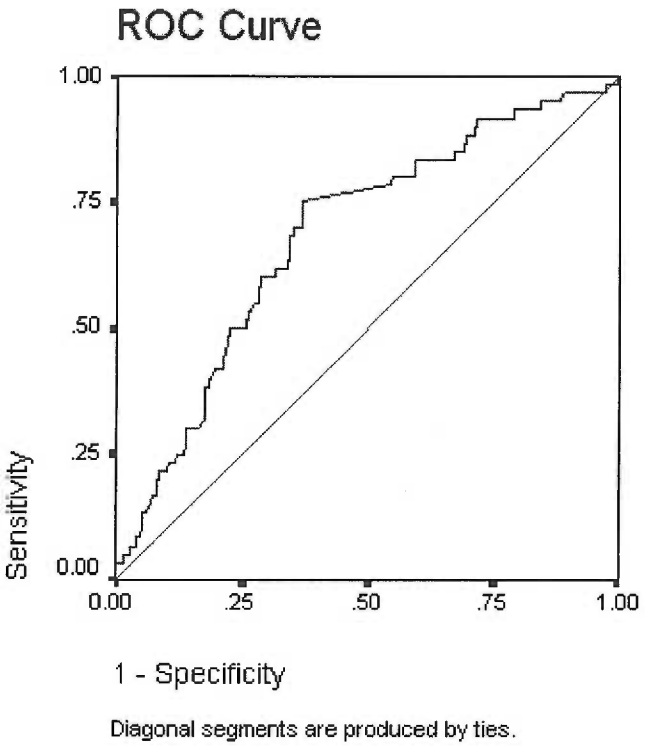


Figure 8: ROC Curve for PSA-categorized with Volume



Appendix A: Figures 2-10
(Continued)

Figure 9: ROC Curve for PSA Density

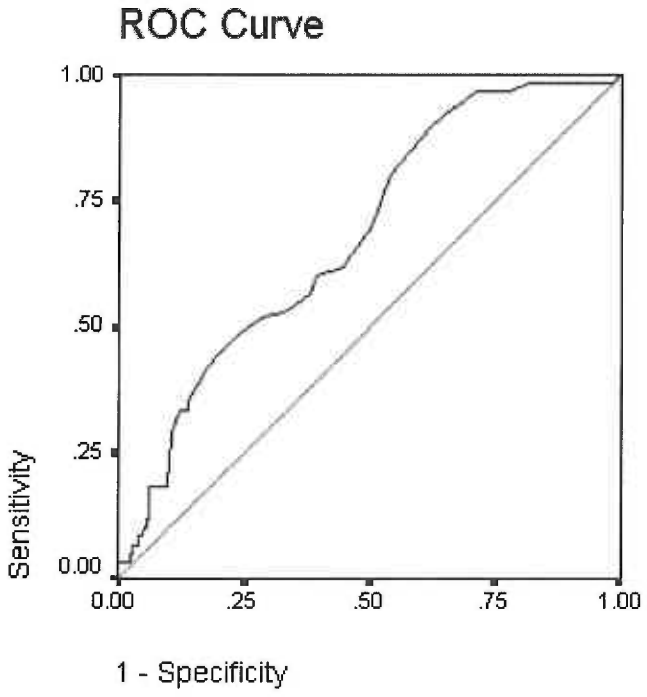


Figure 10: Survival Curves for the Number of Prostate Biopsies (Time in months)