Conditional Probability of Reclassification in an Active Surveillance Program for Prostate Cancer

Ridwan Alam, **H. Ballentine Carter**[*], **Patricia Landis**, **Jonathan I. Epstein**[†], and **Mufaddal Mamawala**[‡]

James Buchanan Brady Urological Institute, The Johns Hopkins University School of Medicine, Baltimore, Maryland

Abstract

Purpose—We evaluated the risk of prostate cancer reclassification by time on active surveillance.

Materials and Methods—From 1995 to 2014 we evaluated 557 and 251 men at very low and at low risk, respectively, who were on active surveillance and compliant with prostate biopsies. Our primary study outcome was reclassification to higher risk disease by grade or extent. Freedom from reclassification was estimated using the Kaplan-Meier approach with adjustment for covariates using the Cox proportional hazards model.

Results—Within the first 2 years of surveillance patient survival free of reclassification by grade (p = 0.20) and by any biopsy criteria (p = 0.25) was similar in men with very low and low risk disease. After 2 years men with low risk disease were 2.4 times more likely to be diagnosed with a Gleason score of greater than 6 than men with very low risk disease (p = 0.002, HR 2.4, 95% CI 1.9–3.5). Additionally, beyond 2 years on surveillance the risk of lifetime reclassification by grade and by any criteria decreased by 30% and 35% (each p <0.0001, HR 0.70, 95% CI 0.60–0.76 and HR 0.65, 95% CI 0.57–0.72, respectively) with each biopsy that showed no reclassification.

Conclusions—The reclassification rate during surveillance is not equally distributed across time or risk groups. Due to misclassification at diagnosis the reclassification rate in very low and low risk groups is similar in the first 2 years but differs significantly beyond 2 years. The risk of reclassification decreases with time for each nonreclassifying biopsy beyond 2 years.

Keywords

prostatic neoplasms; prostate-specific antigen; risk; classification; diagnosis

There is ongoing concern that PSA measurement has led to the diagnosis and treatment of cancer that poses no threat to life.[1,2] Active surveillance with curative intent is an approach to decreasing the harms (overtreatment) of PSA based screening. As an alternative to

[‡]Correspondence: Park 200B, Johns Hopkins Hospital, 600 North Wolfe St., Baltimore, Maryland 21287 (telephone: 410-614-3724; FAX: 410-955-0833; mmamawa1@jhmi.edu).

immediate treatment, active surveillance allows for monitoring of patients at favorable risk with selective delayed intervention in those with disease reclassification (ie higher grade and/or more extensive disease on biopsy). The primary concerns with active surveillance are twofold. 1) Disease may be misclassified at diagnosis so that a patient with aggressive cancer is incorrectly assumed to have low grade cancer. By the time that reclassification occurs upon followup biopsy, perhaps years later, the window of cure is lost.[3,4] 2) Favorable risk cancer may evolve into a more aggressive phenotype during surveillance, closing the window of curability.[3]

Various prognostic factors have been evaluated to determine the risk of reclassification to higher grade cancer in patients on active surveillance, including race, age, biopsy findings, PSA kinetics and imaging.[5,6] However, to our knowledge the interval without reclassification while on active surveillance has not been evaluated as a predictor of future reclassification.

We evaluated the risk of reclassification while on active surveillance as a function of repeat biopsies without reclassification to determine whether absent reclassification could be used to inform patients about the future risk of a higher grade or more extensive prostate cancer diagnosis.

METHODS

Study Cohort

Since January 1995, older men who present to our institution with very low or low risk prostate cancer have been counseled that active surveillance is an acceptable and often preferable alternative to immediate intervention.[7,8] With approval from the institutional review board and informed patient consent we enrolled eligible patients into The Johns Hopkins Active Surveillance Study, an open enrollment, longitudinal study of the natural history of screen detected prostate cancer.[7] A total of 1,298 men were enrolled in the study as of June 2014 and classified as at very low risk (924) or low risk (374) based on the Epstein criteria and D'Amico risk groups.[9,10] As previously described by Epstein et al,[9] the criteria to meet very low risk disease include 1) clinical stage T1c disease, 2) PSAD less than 0.15 ng/ml/cc, 3) Gleason score 6 or less, 4) 2 or fewer positive biopsy cores and 5) a maximum of 50% cancer involvement of any core. Patients who met all 5 Epstein criteria were considered at very low risk even if PSA was 10 ng/ml or greater provided that PSAD was less than 0.15 ng/ml/cc. Patients who did not meet all Epstein criteria but had a Gleason score of 6 or less, PSA less than 10 ng/ml and stage T1c or T2a disease were considered at low risk.[10]

We monitored patients by semiannually measuring PSA with digital rectal examination and annually performing 12-core or greater surveillance biopsy. Curative therapy was discussed based on disease reclassification, defined in our program as upgrading to Gleason score 7 or greater, more than 2 positive cores or greater than 50% cancer involvement of any core. We did not use serum PSA or PSA kinetics as indications for reclassification.

To compare the risk of reclassification among patients at different points in followup we considered the total number of repeat biopsies without reclassification after diagnosis as a proxy for time on active surveillance. We only included patients who were compliant with annual biopsies to ensure the validity of this relationship and that men were at equal risk of reclassification with time. Men were considered compliant if the interval between repeat biopsies was no more than 18 months. Men with more than an 18-month interval between the last biopsy and the last observation date were excluded from study. We identified 557 and 251 at very low and low risk, respectively, patients who fit these criteria for a total of 808 compliant patients in our study. Compared to included men those who were excluded were older (p = 0.01) and had lower PSAD at diagnosis (p = 0.03) and fewer positive cores at diagnosis (p = 0.04). There was also a higher proportion of men at very low risk compared to included patients (p = 0.02). Despite this potential bias these men were excluded to ensure that study patients had a similar probability of reclassifying biopsy with followup. Because preliminary analysis showed that median time to reclassification was about 2 years, the compliant cohort was further dichotomized into patients with less than 2 vs 2 or more years of followup.

Statistical Analysis

Study group characteristics at diagnosis (age, PSA, PSAD, percent free PSA, number of positive biopsy cores, maximum percent involvement of any core with cancer), number of repeat biopsies without reclassification after diagnosis, time to the first surveillance biopsy and time between subsequent surveillance biopsies were evaluated in the compliant cohort and in the entire study cohort. Men reclassified while on active surveillance were further separated into 1 of 3 groups based on time of reclassification, including 1) within the first 2 years of followup, 2) between 2 and 5 years of followup, and 3) after 5 years of followup. Two years represented the median time to reclassification. We used 5 years to describe a long-term distribution of the data.

To ensure validity a bivariate correlation was calculated between followup and the total number of repeat biopsies without reclassification. The annual risk of biopsy reclassification based on Gleason score (upgraded to 7 or greater) and any biopsy criteria (upgraded to Gleason score 7 or greater, more than 2 positive cores or greater than 50% cancer involvement of any core) was calculated as a percent. The 2 reclassification definitions were used for analysis because of differences in the application of this term at various institutions.

We used Kaplan-Meier estimates to track and analyze patients without reclassification while on active surveillance. Time zero was defined as the time of diagnosis and the event of interest was defined as reclassification based on Gleason score or any biopsy criteria. Time to event was calculated in years between time zero and time to reclassification. Men who were not reclassified at the time of analysis, withdrew from the active surveillance program or were lost to followup were censored at the date of the last surveillance visit. Men who died while on active surveillance were censored at time of death. Followup is reported in years between time zero and the time to patient withdrawal, loss to followup or death.

A multivariate Cox proportional hazards model with time dependent covariates was applied to analyze the relationship between patients free of reclassification and age, risk group (very

low vs low), total PSA at diagnosis, PSAD at diagnosis and the number of repeat biopsies without reclassification. Survival free from reclassification was compared in the compliant very low and compliant low risk groups using log likelihood estimates for patients with less than 2 years and 2 or more years of followup.

Statistical significance was considered at p <0.05 for all analyses. Statistical analysis was done with SAS®, version 9.4 and Stata®, version 13.1.

RESULTS

Table 1 lists group characteristics of the 808 compliant patients and the 490 who were excluded from study. Of the 557 men at very low risk 14 (2.5%) died of a cause other than prostate cancer and 54 (9.7%) withdrew or were lost to followup. A total of 126 men (22.6%) were reclassified by Gleason score and 273 men (49.0%) were reclassified by any biopsy criteria. At study end 245 men (44.0%) with very low risk disease were active in the program.

Eight of the 251 patients (3.1%) at low risk died of a cause other than prostate cancer and 10 (4.0%) withdrew or were lost to followup. Reclassification was based on Gleason score and on any biopsy criteria in 109 (43.4%) and 136 men (54.1%), respectively. At the end of the study 128 men with low risk disease (50.9%) were active in the program.

Median followup in the very low and low risk groups was 4.0 (range 0.0 to 18.0) and 3.0 years (0.0 to 18.0), respectively. There were 242 and 55 compliant patients with more than 5 and more than 10 years of followup, respectively. The annual risk of Gleason score reclassification was 4.8% and 11.3% in patients at very low and low risk, respectively. The annual risk of reclassification by any biopsy criterion was 10.3% and 14.1% in very low and low risk cases, respectively. The bivariate correlation between followup and the number of repeat biopsies without reclassification was highly significant (p <0.0001).

Of patients reclassified based on Gleason score 57.1% at very low and 63.3% at low risk were reclassified within the first 2 years (table 2). When reclassification was determined by any biopsy criteria in men who were reclassified, 60.0% at very low and 61.0% at low risk were reclassified within the first 2 years (table 2). A constant denominator was used in each risk group for a given definition of reclassification to determine that the proportions of reclassified patients were not equally distributed across time.

During the first 2 years on active surveillance there was no statistically significant difference between the very low and low risk groups in the risk of Gleason score reclassification or of any biopsy criteria reclassification (p = 0.20 and 0.25, respectively, fig. 1). Figure 1 shows Kaplan-Meier curves of the probability of reclassification in compliant patients with less than 2 years of followup and not the distribution of reclassified patients with time (table 2). Beyond these initial 2 years there was a statistically significantly higher risk of reclassification in men with low risk disease compared to those with very low risk disease (fig. 2). Based on a Cox proportional hazards model men with low risk disease were 2.4 times more likely to be reclassified by Gleason score than their very low risk counterparts (p = 0.002, HR 2.4, 95% CI 1.9–3.5, fig. 2, A). When reclassification was based on any biopsy

criteria, men with low risk disease were 1.8 times more likely to be reclassified than men with very low risk disease (p = 0.014, HR 1.8, 95% CI 1.1–2.9, fig. 2, B).

After adjusting for covariates there was a statistically significant relationship between the risk of reclassification and the number of repeat biopsies in patients with more than 2 years of followup. Beyond the initial 2 years on surveillance men with very low and low risk disease who were not reclassified had a 30% decreased risk of Gleason score reclassification with each subsequent biopsy (p <0.0001, HR 0.70, 95% CI 0.60–0.76, fig. 3, A). Men with very low and low risk disease who had not been reclassified similarly showed a 35% decrease in the risk of any biopsy criteria reclassification with each subsequent biopsy after the first 2 years (p <0.0001, HR 0.65, 95% CI 0.57–0.72, fig. 3, B).

DISCUSSION

Active surveillance of prostate cancer has become an accepted alternative to immediate intervention in patients with favorable risk prostate cancer. However, there is still concern about the risk of cancer progression or misclassification due to biopsy sampling error. We found that whether at very low or low risk most men on active surveillance were reclassified within the first 2 years, most likely due to initial biopsy misclassification. However, after 2 years of surveillance there was a statistically significantly higher risk of reclassification by Gleason score and by any biopsy criteria in men with low risk disease compared to that in men with very low risk disease.

These findings are consistent with those of Tosoian et al.[11] They observed that men with low risk prostate cancer were at twofold greater risk for upgrading and for nonorgan confined cancer at radical prostatectomy compared to men with very low risk prostate cancer. To our knowledge our finding that the risk of reclassification decreased by approximately 30% with each biopsy that did not show reclassification after 2 years on surveillance is new and could inform decisions on the need for surveillance biopsies.

Our overall reclassification rate is consistent with the findings of others, which showed a reclassification rate of 31.0% to 67.4% depending on reclassification criteria.[12–17] Steinberg et al reported that prostate cancer misclassification at diagnosis is most likely due to the sampling error inherent to biopsy.[18] There remains controversy as to how many tissue cores are optimal to minimize misclassification.[19–21] Because most men in our series underwent 12 or 14-core biopsies only, misclassification at diagnosis could explain the high reclassification rate within the first 2 years.

The current study has clinical implications with respect to informing men who are enrolled on active surveillance. The dropout rate from active surveillance in men without evidence of disease reclassification is currently 3% to 18% at various institutions.[4,7,22–24] One potential reason among many may be overriding concern that biopsy does not reflect the seriousness of cancer and, thus, the patient may lose the window of opportunity for cure after reclassification to higher grade.[22,25] Our study shows that men in an active surveillance program can be reassured that reclassification to a higher grade decreases with time after

each biopsy that fails to reveal reclassification beyond 2 years (fig. 3, A). This could motivate men to remain committed to surveillance.

Furthermore, these findings may inform the followup of men on active surveillance protocols. Currently there are no universal parameters to monitor men on active surveillance. Based on our results compliant patients who have been in the program for an extended period without reclassification may be safely monitored with an increased interval between biopsies. Our findings together with improvements in imaging such as multi-parametric magnetic resonance imaging may lead to increased acceptance of active surveillance as a management strategy for prostate cancer.

Our study has inherent limitations that should be considered. Our study cohort represents a select group of patients who were compliant with annual biopsies while on active surveillance. Thus, our results may not be generalizable to other active surveillance populations with less stringent followup criteria. In addition, as a single institution study our sample size was limited.

CONCLUSIONS

We evaluated the risk of reclassification in patients on active surveillance for prostate cancer as a function of the number of repeat surveillance biopsies. The risk of reclassification was not equally distributed across time or risk groups. The reclassification rate was highest in the initial 2 years of surveillance and beyond 2 years it was higher in men with low risk disease than in those with very low risk disease. Furthermore, after 2 years the reclassification risk decreased substantially with each prostate biopsy that did not result in reclassification. Men on active surveillance who are compliant with surveillance biopsies can be reassured that the reclassification risk decreases with each nonreclassifying biopsy beyond 2 years.

Abbreviations and Acronyms

PSA   prostate specific antigen

PSAD   PSA density

References

1. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst. 2003; 95:868. [PubMed: 12813170]

2. Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement Preventive Services Task Force. Ann Intern Med. 2012; 157:120. [PubMed: 22801674]

3. Harlan SR, Cooperberg MR, Elkin E, et al. Time trends and characteristics of men choosing watchful waiting for initial treatment of localized prostate cancer: results from CaPSURE. J Urol. 2003; 170:1804. [PubMed: 14532780]

4. Klotz L, Zhang L, Lam A, et al. Clinical results of long-term follow-up of a large, active surveillance cohort with localized prostate cancer. J Clin Oncol. 2010; 28:126. [PubMed: 19917860]

5. Klotz L, Emberton M. Management of low risk prostate cancer: active surveillance and focal therapy. Curr Opin Urol. 2014; 24:270. [PubMed: 24710055]

6. Fradet V, Kurhanewicz J, Cowan JE, et al. Prostate cancer managed with active surveillance: role of anatomic MR imaging and MR spectroscopic imaging. Radiology. 2010; 256:176. [PubMed: 20505068]

7. Tosoian JJ, Trock BJ, Landis P, et al. Active surveillance program for prostate cancer: an update of the Johns Hopkins experience. J Clin Oncol. 2011; 29:2185. [PubMed: 21464416]

8. Carter HB, Walsh PC, Landis P, et al. Expectant management of nonpalpable prostate cancer with curative intent: preliminary results. J Urol. 2002; 167:1231. [PubMed: 11832703]

9. Epstein JI, Walsh PC, Carmichael M, et al. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. JAMA. 1994; 271:368. [PubMed: 7506797]

10. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA. 1998; 280:969. [PubMed: 9749478]

11. Tosoian JJ, JohnBull E, Trock BJ, et al. Pathological outcomes in men with low risk and very low risk prostate cancer: implications on the practice of active surveillance. J Urol. 2013; 190:1218. [PubMed: 23643603]

12. Davies JD, Aghazadeh MA, Phillips S, et al. Prostate size as a predictor of Gleason score upgrading in patients with low risk prostate cancer. J Urol. 2011; 186:2221. [PubMed: 22014803]

13. Epstein JI, Feng Z, Trock BJ, et al. Upgrading and downgrading of prostate cancer from biopsy to radical prostatectomy: incidence and predictive factors using the modified Gleason grading system and factoring in tertiary grades. Eur Urol. 2012; 61:1019. [PubMed: 22336380]

14. Kvale R, Moller B, Wahlqvist R, et al. Concordance between Gleason scores of needle biopsies and radical prostatectomy specimens: a population-based study. BJU Int. 2009; 103:1647. [PubMed: 19154461]

15. Corcoran NM, Hong MK, Casey RG, et al. Upgrade in Gleason score between prostate biopsies and pathology following radical prostatectomy significantly impacts upon the risk of biochemical recurrence. BJU Int. 2011; 108:E202. [PubMed: 21443656]

16. Shipitsin M, Small C, Choudhury S, et al. Identification of proteomic biomarkers predicting prostate cancer aggressiveness and lethality despite biopsy-sampling error. Br J Cancer. 2014; 111:1201. [PubMed: 25032733]

17. Corcoran NM, Hovens CM, Hong MK, et al. Underestimation of Gleason score at prostate biopsy reflects sampling error in lower volume tumours. BJU Int. 2012; 109:660. [PubMed: 21895937]

18. Steinberg DM, Sauvageot J, Piantadosi S, et al. Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. Am J Surg Pathol. 1997; 21:566. [PubMed: 9158682]

19. Eichler K, Hempel S, Wilby J, et al. Diagnostic value of systematic biopsy methods in the investigation of prostate cancer: a systematic review. J Urol. 2006; 175:1605. [PubMed: 16600713]

20. Delongchamps NB, de la Roza G, Jones R, et al. Saturation biopsies on autopsied prostates for detecting and characterizing prostate cancer. BJU Int. 2009; 103:49. [PubMed: 18680497]

21. Capitanio U, Karakiewicz PI, Valiquette L, et al. Biopsy core number represents one of foremost predictors of clinically significant Gleason sum upgrading in patients with low-risk prostate cancer. Urology. 2009; 73:1087. [PubMed: 19195695]

22. Berger ZD, Yeh JC, Carter HB, et al. Characteristics and experiences of patients with localized prostate cancer who left an active surveillance program. Patient. 2014; 7:427. [PubMed: 24920082]

23. Cooperberg MR, Carroll PR, Klotz L. Active surveillance for prostate cancer: progress and promise. J Clin Oncol. 2011; 29:3669. [PubMed: 21825257]

24. Soloway MS, Soloway CT, Williams S, et al. Active surveillance; a reasonable management alternative for patients with prostate cancer: the Miami experience. BJU Int. 2008; 101:165. [PubMed: 17850361]

25. Carter HB, Partin AW, Walsh PC, et al. Gleason score 6 adenocarcinoma: should it be labeled as cancer? J Clin Oncol. 2012; 30:4294. [PubMed: 23032616]
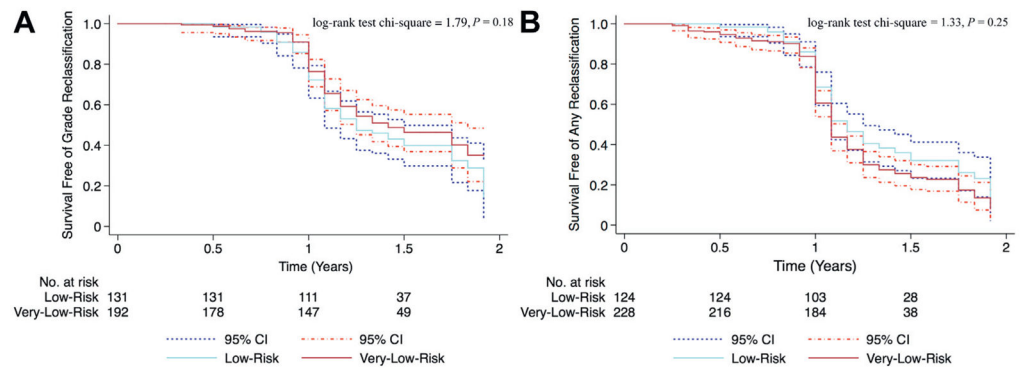
Figure 1.
Kaplan-Meier curves of reclassification in compliant patients (not all patients on active surveillance) with less than 2-year followup. *A*, Gleason score. *B*, any biopsy criteria.
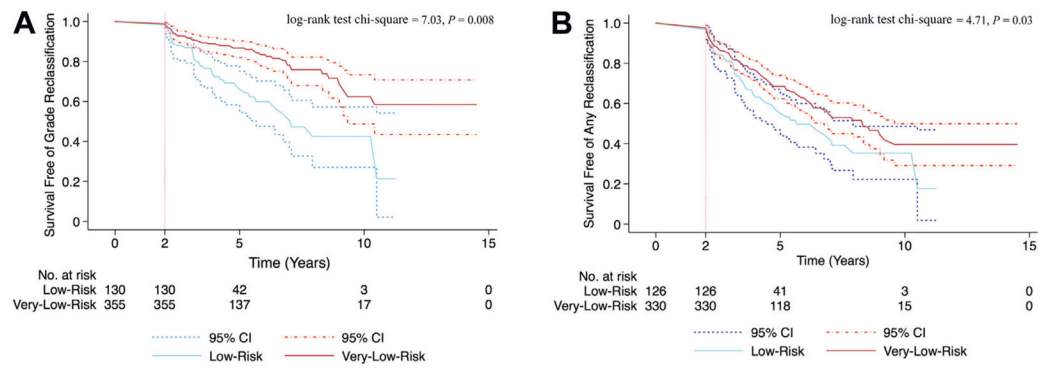
Figure 2.
Kaplan-Meier curves of reclassification in compliant patients (not all patients on active surveillance) not reclassified within first 2 years. *A*, Gleason score. *B*, any biopsy criteria.
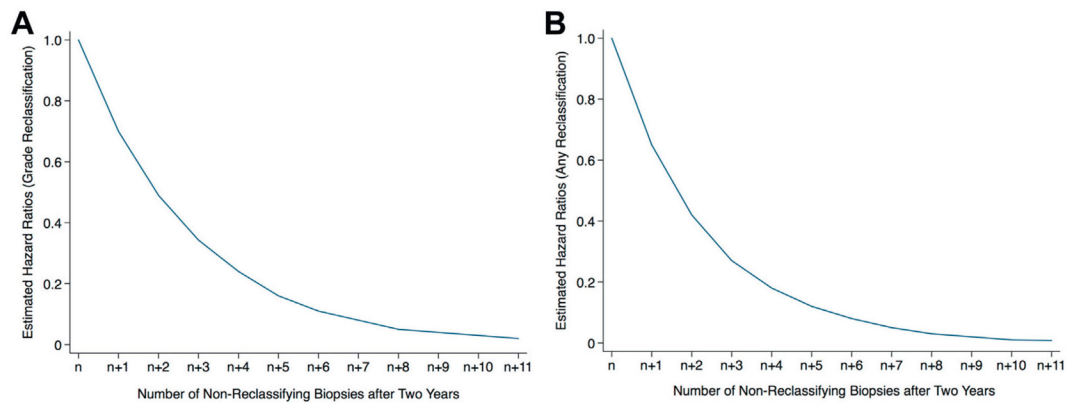
Figure 3.
Estimated risk of reclassification in compliant patients after 2 years of active surveillance. Number of biopsies at end of 2 years varied even in compliant patients based on interval between biopsies. *A*, Gleason score. *B*, any biopsy criteria.

Table 1

Demographics of 808 study and 490 excluded patients

	Mean ± SD Study Group/Median (range)	Mean ± SD Excluded Group/Median (range)	p Value
At diagnosis:			
Age	65.4 ± 5.7/66 (41–83)	65.7 ± 6.1/66 (41–92)	0.01
PSA (ng/ml)	5.2 ± 2.6/4.8 (0.18–24)	5.2 ± 2.8/4.8 (0.18–24)	0.40
PSAD (ng/ml/cc)	0.16± 0.06/0.10 (0.02–0.69)	0.11± 0.06/0.10 (0.004–0.69)	0.03
% Free PSA[*]	19.3 ± 7.2/18.0 (5–44)	18.2 ± 7.3/17.2 (4–50)	0.30
No. pos cores	1.4 ± 0.7/1.0 (1–7)	1.3 ± 0.7/1.0 (1–7)	0.04
Max % any pos core involvement	10.6 ± 14.7/5.0 (1–100)	10.0 ± 14.7/5.0 (1–100)	0.20
Time to 1st surveillance biopsy (mos)	11.0 ± 3.7/12.0 (1–18)	12.4 ± 8.2/12.0 (1–91)	<0.0001
Time between subsequent surveillance biopsies (mos)	12.6 ± 2.0/12.0 (2–18)	14.5 ± 6.8/12.0 (0–131)	<0.0001
No. repeat biopsies after diagnosis until reclassification	3 ± 2.3/2 (0–11)	3 ± 2.0/2 (0–12)	0.60

[*] Not measured in all patients.

Table 2

Distribution of reclassified patients on active surveillance with time

Yrs to Reclassification	No. Compliant Very Low Risk Pts (%)	No. Compliant Low Risk Pts (%)
Gleason score:	126	109
Less than 2	72 (57.1)	69 (63.3)
2–5	37 (29.4)	31 (28.4)
Greater than 5	17 (13.5)	9 (8.3)
Any biopsy criteria:	273	136
Less than 2	164 (60.0)	83 (61.0)
2–5	83 (30.4)	44 (32.3)
Greater than 5	26 (9.6)	9 (6.6)