Pathologic Outcomes in Favorable-risk Prostate Cancer: Comparative Analysis of Men Electing Active Surveillance and Immediate Surgery

Jeffrey J. Tosoian a, Debasish Sundi a, Bruce J. Trock a, b, c, Patricia Landis a, Jonathan I. Epstein a, c, d, Edward M. Schaeffer a, c, d, H. Ballentine Carter a, c, Mufaddal Mamawala a, *

* The James Buchanan Brady Urological Institute, Johns Hopkins Medical Institutions, Baltimore, MD, USA; b Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; c Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD, USA; d Department of Pathology, Johns Hopkins Medical Institutions, Baltimore, MD, USA

Article history:
Accepted September 21, 2015

Associate Editor:
James Catto

Keywords:
Active surveillance
Oncologic outcomes
Prostate cancer

Abstract

Background: It remains unclear whether men selecting active surveillance (AS) are at increased risk of unfavorable longer term outcomes as compared with men who undergo immediate treatment.

Objective: To compare adverse pathologic outcomes in men with favorable-risk prostate cancer who underwent delayed prostatectomy after surveillance (DPAS) to those who elected immediate prostatectomy (IRP).

Design, setting, and participants: We conducted a retrospective analysis of a prospective AS registry from 2004 to 2014. From the Johns Hopkins AS program (n = 1298), we identified a subset of men who underwent DPAS (n = 89) and was representative of the entire cohort, not just those that were reclassified to higher risk. These men were compared with men who underwent IRP (n = 3788).

Outcome measurements and statistical analysis: We measured adverse pathologic features (primary Gleason pattern ≥4, seminal vesicle invasion [SVI], or lymph node [LN] positivity). Multivariable models were adjusted for age, prostate-specific antigen density, and baseline risk classification.

Results and limitations: Delayed prostatectomy occurred at a median of 2.0 yr (range: 0.6–9.0) after diagnosis. The DPAS and IRP cohorts demonstrated similar proportions of men with primary Gleason pattern ≥4 (17% vs 20%; p = 0.11), SVI (3.3% vs 3.2%; p = 0.53), LN positivity (2.3% vs 1.2%; p = 0.37), and overall adverse pathologic features (21.3% vs 17.0%; p = 0.32). The adjusted odds ratio of adverse pathology was 1.33 (95% confidence interval, 0.82–2.79; p = 0.13) for DPAS as compared with IRP. Limitations include a modest cohort size and a limited number of events.

Conclusions: In men with favorable-risk cancer, the decision to undergo AS is not independently associated with adverse pathologic outcomes.

Patient summary: This report compares men with favorable-risk prostate cancer who elected active surveillance with those who underwent immediate surgery accounting for evidence that approximately one-third of men who choose surveillance will eventually undergo treatment. Our findings suggest that men who are closely followed with surveillance may have similar outcomes to men who elect immediate surgery, but additional research is needed.

© 2015 European Association of Urology. Published by Elsevier B.V. All rights reserved.
1. Introduction

Curative intervention of prostate cancer (PCa) is associated with side effects that are likely to have an impact on quality of life [1,2]. Although morbidity may be accepted in the setting of life-extending therapy, the treatment of indolent nonlethal tumors exposes men to unnecessary negative effects [3]. Despite improvements in diagnostic and imaging techniques, it remains difficult to predict which cancers will and will not prove harmful during a man’s lifetime [4]. Active surveillance (AS) has been proposed as a potential solution to this problem [5].

Although an increasing proportion of men are being managed with AS [6,7], variability in utilizing this approach is large and depends more on physician practice patterns than tumor metrics [8]. One factor contributing to the limited uptake of AS is a fear that failure to treat a localized cancer while in a curable stage could prove costly over a long-term follow-up [9]. The vast majority of favorable-risk cancers demonstrate a prolonged, indolent course [10], but approximately a third will eventually demonstrate higher risk features [11]. In the absence of prospective trials comparing various management strategies [12], previous studies have compared surgical outcomes of men undergoing immediate radical prostatectomy (IRP) with subjects initially managed on AS who subsequently underwent delayed prostatectomy after an interval on surveillance [13–15].

A major limitation of this approach, however, is that most men on AS who proceed to surgery do so based on worrisome findings such as biopsy reclassification or increasing prostate-specific antigen (PSA) [15,16]. Thus delayed radical prostatectomy (RP) cohorts have been composed of the highest risk AS patients who ultimately failed this approach. Although informative to this subgroup of patients who fail AS during follow-up, such comparisons provide minimal insight to the newly diagnosed man who does not know whether he will later exhibit high-risk disease. As such, others have proposed that an ideal study design would compare an immediate surgery population with a similar cohort that selected AS at the same time—including both the minority of AS patients who progressed to treatment and the majority of AS patients who remained on surveillance [14]. Because this third group does not undergo treatment, however, treatment outcomes are not available for assessment. By incorporating a population of men who elected to undergo treatment in the absence of progression, however, we uniquely compared pathologic outcomes from a risk-representative delayed surgery after surveillance cohort with a similar cohort that underwent IRP.

2. Methods

2.1. Active surveillance cohort

Starting in 1995, men with favorable-risk (ie, very low risk [VLR] or low risk [LR]) PCa were offered enrollment in AS [17]. VLR criteria include clinical stage T1c disease, prostate-specific antigen density (PSAD) <0.15 ng/ml, biopsy Gleason score (GS) ≤6, two or fewer positive biopsy cores, and ≤50% involvement of any core with cancer. We have not used PSA (eg, >10 ng/ml) to exclude men from VLR classification if PSAD was <0.15 because our experience has not identified a clinically meaningful PSA cut-off to predict higher risk disease [18]. LR cancer was defined as clinical stage ≤T2a, PSA <10 ng/ml, and GS ≤6. Our protocol includes semiannual PSA and digital rectal examination, and annual prostate biopsy in most cases. As previously described, adherence to our protocol approximates 90% annually [19]. Curative intervention is recommended upon disease reclassification, defined as biopsy findings no longer meeting inclusion criteria. Based on our experience, PSA kinetics are not used as a trigger for reclassification in this program [20]. As such, interventions not due to biopsy reclassification are due to changes in patient preference.

Of 1298 men enrolled in AS during the study period, 926 men (71.3%) met VLR and 372 (28.7%) met LR classification criteria. During follow-up (median: 5.0 yr; range: 0.01–18.0), 467 men (36.0%) underwent reclassification at a median of 2.0 yr after diagnosis; conversely, 831 men (64% of the overall cohort) did not undergo reclassification and were eligible to remain on AS. Of those who reclassified, 81 (6.5%) reclassified based on grade, 234 (18.0%) reclassified based on volume, and 149 (11.5%) reclassified based on grade and volume.

2.2. Study design

We sought to assess an intermediate end point (ie, surgical pathology) in the AS population for comparison with men who underwent immediate treatment. Comparing the entire delayed RP population, however, with men who undergo IRP is inherently biased against AS; the delayed RP cohort does not represent all men who select AS, but rather those men who select AS and then demonstrate higher risk disease during follow-up (Fig. 1). This bias has been described in the setting of AS by multiple authors and is a well-recognized limitation of previous comparative studies [14]. To mitigate this limitation, we aimed to identify a delayed prostatectomy after surveillance (DPAS) population that represented the overall AS cohort in terms of demographic factors, reclassification rate during follow-up (36%), and type of reclassification observed (6.5% by grade, 18% by volume, and 11.5% by grade and volume). To ensure modern pathologic grading, the study was limited to men who underwent surgery after 2004.

2.3. Patient population

2.3.1. Delayed prostatectomy after surveillance (study cohort)

From January 2005 to September 2014, a total of 185 men underwent delayed prostatectomy after initial management on AS. Sixty-one men elected to undergo RP in the absence of clinical or pathologic evidence of disease progression. Four of these men (6.6%) were excluded due to age (median age: 45.5 yr; range: 44–46 yr) discordant with the AS population, yielding 57 study-eligible men who underwent delayed RP without a trigger for intervention. Consistent with our AS experience, these 57 men were designated to represent 64% of the DPAS study cohort, and the total DPAS cohort was calculated to equal 89 subjects. The study design is illustrated in Figure 2.

Corresponding to reclassification event rates in the overall AS cohort, the remaining 36% (n = 32) of the DPAS study group was composed of 6 men (6.7%) who underwent grade reclassification, 16 (18%) who underwent volume reclassification, and 10 (11.3%) who underwent grade plus volume reclassification. To minimize selection bias, subjects were selected from eligible men within each reclassification group using simple random sampling after stratification by year of surgery. Patient characteristics were similar between the DPAS cohort and overall AS cohort, confirming representative sampling (Supplementary Table 1).
2.3.2 Immediate prostatectomy (control cohort)

The IRP comparison group was composed of 3788 men who underwent immediate prostatectomy for favorable-risk disease as diagnosed by a 12-core biopsy; these data were obtained from a prospectively collected institutional database [21]. IRP was defined as within 1 yr of diagnosis, although most patients underwent IRP within 4 mo (median time to surgery: 3.0 mo; interquartile range [IQR]: 2.0–4.0 mo).

2.4 Outcomes

The primary end points were pathologic features after RP including GS, pathologic tumor stage, positive surgical margins, and lymph node (LN) positivity. Adverse pathology was considered as a composite outcome (ie, primary Gleason pattern 4 or 5, seminal vesicle invasion [SVI], or LN positivity) based on the association of this outcome with 15-yr PCa-specific mortality [22]. LN dissection is a provider-specific intervention at our institution; dissection was performed in 91% of patients assessed in this study cohort. Biochemical recurrence (BCR) was a secondary outcome defined as two consecutive PSA increases >0.2 ng/ml.

2.5 Statistical analysis

The DPAS and IRP cohorts were compared using the t test, Cochran t test, and chi-square analysis as appropriate. Unconditional multivariable logistic regression was used to compare outcomes based on initial treatment choice (ie, AS vs IRP). Factors significant on univariable analysis were entered into a multivariable model to control for potential confounding effects.
analysis were included in the multivariable model. To ensure stable and valid results, we chose to limit the number of input variables by using VLR/LR status in place of collinear variables used to define risk status (ie, number of positive cores, maximum percentage involvement of any core) [23]. The resultant odds of adverse pathology were unchanged in a sensitivity analysis including the distinct variables (odds ratio [OR]: 1.41; 95% confidence interval [CI], 0.73–2.66). BCR was assessed on survival analysis and compared using the log-rank test. Post hoc power estimation was performed using the Wald test, adjusting for confounders in the logistic model. All statistical analyses were performed using SAS v.9.2 (SAS Institute, Cary, NC, USA) and were two sided; \( p < 0.05 \) was considered statistically significant.

### 3. Results

Men in the DPAS study cohort underwent RP at a median of 2.0 yr after diagnosis (IQR: 1.0–4.0 yr) as compared with 3.0 mo (IQR: 2.0–4.0 mo) in the comparison (IRP) cohort. As demonstrated in Table 1, the DPAS cohort was older (\( p < 0.001 \)) and had higher PSAD (\( p = 0.04 \)) than the IRP cohort. The DPAS cohort had a lower mean number of cores positive for cancer (1.5 vs 2.9; \( p < 0.001 \)) and lower mean percentage core involvement (8.3 vs 35.3; \( p < 0.001 \)). Accordingly, the DPAS cohort had a greater proportion of men with VLR disease (70.8% vs 43.1%; \( p < 0.001 \)).

Table 2 shows the pathologic outcomes. The DPAS cohort demonstrated a nonsignificant trend toward higher GS (\( p = 0.26 \)). Positive surgical margins were observed in 12 men (13.5%) in the DPAS cohort and 439 (11.6%) in the IRP cohort (\( p = 0.50 \)); a positive LN was observed in 2.3% of the DPAS cohort versus 1.2% of the IRP cohort (\( p = 0.37 \)). Ultimately, adverse pathology was present in 19 of 89 men (21.3%) in the DPAS cohort and 643 of 3788 (17.0%) in the IRP group (\( p = 0.32 \)). In the multivariable model adjusted for age, PSAD, and preoperative risk classification (ie, VLR/LR), the OR for adverse pathology was 1.33 (95% CI, 0.82–2.79; \( p = 0.13 \)) in the DPAS cohort as compared with the IRP cohort (Table 3). Post hoc power estimation using the Wald test for logistic regression was 0.573.

Median follow-up for assessment of BCR was 2.0 yr in the DPAS cohort and 3.0 yr in the IRP cohort. BCR was observed in 6 men (6.7%) in the DPAS group and 160 (4.2%) in the IRP cohort, and median interval to BCR was not reached in either group. On survival analysis there was no significant difference in time to BCR rate in the groups (\( p = 0.10 \)).

### Table 1 – Demographics of the study (delayed prostatectomy) and control (immediate radical prostatectomy) cohorts

<table>
<thead>
<tr>
<th></th>
<th>DPAS cohort (n = 89)</th>
<th>IRP cohort (n = 3788)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td>Age at diagnosis, yr</td>
<td>65</td>
<td>55–83</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.0001 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.0001 )</td>
<td></td>
</tr>
<tr>
<td>PSA, ng/ml</td>
<td>5.0</td>
<td>0.7–17</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.0001 )</td>
<td></td>
</tr>
<tr>
<td>PSAD</td>
<td>0.1</td>
<td>0.02–0.41</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.0001 )</td>
<td></td>
</tr>
<tr>
<td>Biopsy cores sampled, n</td>
<td>12</td>
<td>6–44</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.0001 )</td>
<td></td>
</tr>
<tr>
<td>Cores positive for cancer, n</td>
<td>1</td>
<td>1–3</td>
</tr>
<tr>
<td></td>
<td>( p &lt; 0.0001 )</td>
<td></td>
</tr>
<tr>
<td>Maximum % involvement of any core with cancer</td>
<td>5</td>
<td>1–30</td>
</tr>
</tbody>
</table>

### Table 2 – Pathologic outcomes by study group

<table>
<thead>
<tr>
<th></th>
<th>DPAS (n = 89)</th>
<th>IRP (n = 3788)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
<td>%</td>
</tr>
<tr>
<td>Gleason score</td>
<td>( 3 + 3 )</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>( 3 + 4 )</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>( 4 + 3 )</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>( &gt;4 + 4 )</td>
<td>4</td>
</tr>
<tr>
<td>Pathologic stage</td>
<td>pT2</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>pT3a</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>pT3b</td>
<td>3</td>
</tr>
<tr>
<td>Positive surgical margins</td>
<td>12</td>
<td>13.5</td>
</tr>
<tr>
<td>Lymph node positive</td>
<td>2</td>
<td>2.3</td>
</tr>
<tr>
<td>Adverse pathology*</td>
<td>19</td>
<td>21.3</td>
</tr>
</tbody>
</table>

### Note

- DPAS = delayed prostatectomy after surveillance; IRP = immediate radical prostatectomy; LR = low risk; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density; SD = standard deviation; VLR = very low risk.
- * Primary Gleason pattern 4 or 5, seminal vesicle invasion, or lymph node positivity.
Table 3 – Multivariable analysis for adverse pathology

<table>
<thead>
<tr>
<th></th>
<th>Univariable analysis</th>
<th>Multivariable analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>p</td>
</tr>
<tr>
<td>Delayed vs immediate</td>
<td>1.24 (0.80–2.05)</td>
<td>0.27</td>
</tr>
<tr>
<td>Age per year</td>
<td>1.06 (1.04–1.08)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>PSAD per 0.1-unit increase</td>
<td>1.22 (1.09–1.44)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>VLR vs LR</td>
<td>0.64 (0.48–0.86)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Ci = confidence interval; LR = low risk; OR = odds ratio; PSAD = prostate-specific antigen density; VLR = very low risk.

4. Discussion

AS is one strategy to mitigate the overtreatment associated with PSA-based screening. The absence of data comparing men on AS with immediate treatment populations has made it difficult to quantify the long-term risks associated with choosing AS. We found that the likelihood of harboring adverse pathology was not significantly different in men with favorable-risk PCa who initially opted for AS versus those who underwent IRP. Acknowledging that our outcome of interest was an intermediate pathologic end point, these findings suggest that the decision to defer IRP for AS may not compromise longer term outcomes in men willing to undergo careful monitoring.

Delayed surgery cohorts in previous comparisons were enriched with men who failed AS during follow-up. Iremashvili et al examined 22 men who underwent DPAS, of whom 16 (73%) demonstrated biopsy GS upgrading prior to surgery and the remaining 6 (27%) underwent volume reclassification [24]. In RP specimens, they observed a greater proportion of primary GS ≥4 (32% vs 11%; p = 0.023) and a higher median percentage of cancer (12.5% vs 5.0%; p = 0.009) in the DPAS cohort as compared with the IRP cohort, although there were similar rates of extraprostatic extension, SVI, LN positivity, and positive margins. Satkunasivam and colleagues assessed a DPAS cohort of 41 men, of whom 68% demonstrated biopsy GS upgrading, 12% demonstrated volume reclassification, and 7% had a rapidly rising PSA level prior to surgery [15]. The authors reported significantly higher likelihood of GS ≥7 (65.9% vs 37.6%; p = 0.004), pT3 disease (36.6% vs 10.7%; p = 0.0002), and positive margins (14.6% vs 1.8%; p = 0.007) in the DPAS cohort.

Certainly most studies suggest at least a trend toward less favorable outcomes in this subgroup. Understanding such risks is essential in counseling this population of men who progress on AS, approximately a third of those enrolling, regarding the probability of adverse outcomes should they proceed to surgery. At the same time, it is essential to recognize that these studies address a question fundamentally different from ours, quantifying the risk posed specifically to those men who enroll in AS and then fail during follow-up. By comparison, our study sought to assess the risk associated with initially choosing AS while factoring in the likelihood that approximately a third of such men will progress to treatment and acknowledging our inability prospectively to identify which men will progress [11]. Facing a decision between immediate treatment and AS, our study provides insight for the newly diagnosed patient who does not know whether he will progress during monitoring and who would like to consider his risk of negative longer term outcomes should he elect AS.

Acknowledging use of an intermediate end point, selecting AS for initial management of favorable-risk disease does not appear to predict unfavorable outcomes. Certainly, as Filipou et al and others have demonstrated [25], clinical findings during AS have a significant impact on longer term outcomes. Given that contemporary techniques cannot accurately predict which men will fail AS during follow-up, the ability to characterize risk to all AS-eligible men should prove helpful in counseling and decision making. Discussing unfavorable outcomes after a trial on AS is a difficult scenario, but evidence indicates that well-informed patients are more engaged, more likely to fully consider the risks and benefits of treatment options, and more satisfied with clinical encounters [26]. These attributes are undeniably critical to shared decision making, and such findings further emphasize the importance of discussing risks and benefits prior to proceeding with any management strategy.

The current study was limited by a relatively small study cohort, and it is possible that the observed trend toward more frequent adverse outcomes in AS would reach statistical significance in a larger population. Nonetheless, this study is the first to our knowledge to use a risk-stratified cohort corresponding to contemporary AS populations instead of only the highest risk men, and our findings should prove useful as a starting point in discussing AS. Because our program utilizes conservative selection and monitoring criteria, these results may differ in cohorts composed of higher risk men and those with less stringent follow-up. Most men were white, and our findings may not apply to more racially diverse cohorts. At the same time, magnetic resonance imaging was not formally included in our protocol during the study period, and use of additional aids such as imaging or genomic testing may have an impact on the applicability of our findings. Finally, although well-validated in multiple cohorts, the intermediate end points used in this study may not completely reflect the outcomes most important to patients such as freedom from PCa death and metastasis.

5. Conclusions

Our findings suggest that the decision to initiate AS of favorable-risk PCa is not independently associated with adverse pathologic findings. Observations during the course
of surveillance certainly influence longer term outcomes, and additional research is needed to better identify patient-level factors most predictive of prognosis.

**Author contributions:** Mufaddal Mamawala had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Tosoian, Sundi, Trock, Epstein, Schaeffer, Carter, Mamawala.

**Acquisition of data:** Mamawala.

**Analysis and interpretation of data:** Tosoian, Sundi, Trock, Carter, Mamawala.

**Drafting of the manuscript:** Tosoian, Sundi, Carter, Mamawala.

**Critical revision of the manuscript for important intellectual content:** Epstein, Schaeffer, Sundi.

**Statistical analysis:** Mamawala.

**Obtaining funding:** None.

**Administrative, technical, or material support:** Landis.

**Supervision:** None.

**Other (specify):** None.

**Financial disclosures:** Mufaddal Mamawala certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

**Funding/Support and role of the sponsor:** None.

**Appendix A. Supplementary data**

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.euro.2015.09.032.

**References**


