Understanding the Performance of Active Surveillance Selection Criteria in Diverse Urology Practices

Scott R. Hawken,* Paul R. Womble,* Lindsey A. Herrel,† Zaojun Ye, Susan M. Linsell, Patrick M. Hurley,‡ James E. Montie§ and David C. Miller‖,¶ for the Michigan Urological Surgery Improvement Collaborative

From the Department of Urology, University of Michigan, Ann Arbor and Providence Hospital (PMH), Southfield, Michigan

Purpose: We used data from MUSIC (Michigan Urological Surgery Improvement Collaborative) to evaluate the performance of published selection criteria for active surveillance in diverse urology practice settings.

Materials and Methods: For several active surveillance guidelines we calculated the proportion of men meeting each set of selection criteria who actually entered active surveillance, defined as the sensitivity of the guideline. After identifying the most sensitive guideline for the entire cohort we compared demographic and tumor characteristics between patients who met this guideline and entered active surveillance, and those who received initial definitive therapy.

Results: Of 4,882 men with newly diagnosed prostate cancer 18% underwent active surveillance. When applied to the entire cohort, the sensitivity of published guidelines ranged from 49% in Toronto to 62% at Johns Hopkins. At a practice level the sensitivity of Johns Hopkins criteria varied widely from 27% to 84% ($p < 0.001$). Compared with men undergoing active surveillance, those meeting Johns Hopkins criteria who received definitive therapy were younger ($p < 0.001$) and more likely to have a positive family history ($p = 0.003$), lower prostate specific antigen ($p < 0.001$), a greater number of positive cores (2 vs 1) on biopsy ($p < 0.001$) and a higher cancer volume in positive core(s) ($p = 0.002$).

Conclusions: The sensitivity of published active surveillance selection criteria varies widely across diverse urology practices. Among patients meeting the most stringent criteria those who received initial definitive therapy had characteristics suggesting greater cancer risk, underscoring the nuanced clinical factors that influence treatment decisions.

Key Words: prostatic neoplasms, watchful waiting, quality improvement, standards, risk

Abbreviations and Acronyms

AS = active surveillance

CCI = Charlson comorbidity index score

GPC = greatest percentage of biopsy core involved with cancer

JH = Johns Hopkins

NCCN® = National Comprehensive Cancer Network®

PSA = prostate specific antigen

Accepted for publication May 5, 2015.

Study received local institutional review board approval.

* Equal study contribution.

† Financial interest and/or other relationship with ArborMetrix.

‡ Financial interest and/or other relationship with Greater Michigan Lithotripsy and AKSM.

§ Financial interest and/or other relationship with HistoSonics.

‖ Correspondence: Department of Urology, University of Michigan, NCRC Building 16, Room 108E, 2800 Plymouth Rd., Ann Arbor, Michigan 48109 (telephone: 734-936-0054; FAX: 734-232-2400; e-mail: dcmiller@med.umich.edu).

¶ Financial interest and/or other relationship with ArborMetrix, Blue Cross Blue Shield of Michigan, MUSIC and Michigan Value Collaboration.

DUE to concerns about overtreatment of men with lower risk prostate cancer[1–4] initial AS is being used more frequently for patients with early stage disease.[5] The potential benefits of AS include avoidance of treatment related side effects (eg urinary incontinence and erectile dysfunction) by delaying or not pursuing definitive therapy. However, these benefits must be weighed against the potential risk of cancer progression. Given the growing acceptance of AS among patients and physicians, there are now many published guidelines describing optimal selection criteria for patients to enter initial AS.[6–13]

0022-5347/15/1945-1253/0
THE JOURNAL OF UROLOGY®

http://dx.doi.org/10.1016/j.juro.2015.05.014
Vol. 194, 1253-1257, November 2015
Printed in U.S.A.

Although most guidelines for entry into initial AS are widely recognized,[6–13] their application and performance in clinical practice are not well characterized. One measure of interest is the sensitivity of such guidelines in the real world setting, that is how many men who meet these selection criteria actually proceed to initial AS. While it makes sense that the most restrictive criteria would also be most sensitive, the actual proportion of patients undergoing surveillance according to these different guidelines is not well defined for diverse urology practice settings. In addition, little is known about the factors that lead men who meet established criteria for entering surveillance to nonetheless proceed to initial local therapy.

In this context we examined the sensitivity of several published guidelines for identifying men who elected initial AS in the community and academic practices comprising MUSIC. In addition to comparing the sensitivity of these guidelines across MUSIC practices, we evaluated differences in demographic characteristics, comorbidity and cancer severity among patients meeting the most sensitive selection criteria who entered AS and those who underwent initial definitive therapy.

MATERIALS AND METHODS

Michigan Urological Surgery Improvement Collaborative

MUSIC was established in 2011 with the aim of improving the quality and cost efficiency of prostate cancer care in Michigan.[5,14–16] The collaborative now includes 42 urology practices comprising more than 90% of urologists in the state. MUSIC receives financial support from Blue Cross® Blue Shield® of Michigan. Each participating practice obtained exemption or approval for participation from a local institutional review board.

Study Population and Data Elements

For all men who undergo prostate biopsy and/or have a new prostate cancer diagnosis seen in participating practices trained abstractors enter a standardized set of data elements in a web based registry, including age, race/ethnicity, CCI, serial PSA results, clinical stage, biopsy Gleason score, number of positive cores, cancer directed treatments, and followup laboratory and pathology results. Quality assurance steps for MUSIC data have been described previously.[5,14–16] The population for this analysis included 4,883 men with newly diagnosed prostate cancer entered in the MUSIC registry from March 2012 through June 2014. Details of participating urologists and practices can be found at www.musicurology.com.

Identification

Primary Treatment. To ensure complete and accurate data, abstractors wait 3 months from the date of prostate cancer diagnosis before entering information about cancer treatment in the MUSIC registry. In addition, explicit documentation in the medical record is required to assign a specific treatment (eg AS, radical prostatectomy, external beam radiation therapy, etc). For instance, the definition of AS in MUSIC (as defined by participating urologists and provided to the data abstractors) is, "Active surveillance is a slightly more structured and aggressive form of watchful waiting. It is recommended that a patient undergo a DRE and a PSA test every 3 or 6 months, depending on the patient's precise history and clinical condition and to re-biopsy the patient yearly or every two years. The physician monitors the patient aggressively and will regularly discuss disease status with the patient so that joint decisions are made about the need for actual treatment." As described by Womble et al, treatment assignment in MUSIC has been externally validated with claims data with excellent concordance.[5]

Active Surveillance Selection Criteria. For this analysis we identified several of the most prominent guidelines for selecting patients for initial AS, including those from certain institutions and organizations (JH,[6] NCCN low and very low risk,[7] Memorial Sloan Kettering Cancer Center,[8] University of California-San Francisco[9,10] and University of Toronto[11,12]). Selection criteria for each of these guidelines are based on routinely available clinical data but vary according to specific elements and definitions.

Statistical Analyses

We first compared clinical and pathological characteristics of patients who entered AS and those who received other initial treatment. We next calculated the proportion of all men meeting each set of selection criteria who entered AS as well as the associated standard Wald asymptotic 95% CIs. We refer to this proportion throughout the study as the sensitivity of real world practice patterns for each guideline. The numerator comprises men who met each selection criterion and chose AS. The denominator comprises all men who met the selection criteria.

After identifying the guideline with the most sensitive selection criteria for the entire study population (ie for the collaborative as a whole), we examined variation in the sensitivity of this guideline at the practice level for sites with greater than 10 patients who met these criteria. We chose to examine the most sensitive guideline to identify the cohort of patients that the greatest number of urologists would likely agree were candidates for surveillance. We then used the chi-square test to examine differences in the sensitivity of this guideline across MUSIC practices.

Finally for men meeting selection criteria for the most sensitive guideline we examined differences between patients who entered AS and those who received definitive therapy. Because we were only interested in men considered eligible to receive local therapy, we excluded from this analysis 47 treated with watchful waiting and 4 treated with androgen deprivation therapy. We analyzed differences in variables that are not included explicitly in the guideline selection criteria (eg age, race, comorbidity and practice size) as well as those that are included in the criteria but still maintained a range of clinically meaningful values after selection, such as PSA level, PSA density, number of positive biopsy cores and GPC. Using the same variables we also compared men who met the most sensitive AS guidelines and received

radiation vs those who underwent surgery. For each comparison we used the 2-sided t-test, the Wilcoxon rank sum test or the chi-square test as appropriate. All statistical testing was performed at the 5% significance level using SAS®, version 9.3.

RESULTS

Table 1 lists the clinical characteristics of the 4,883 men with newly diagnosed prostate cancer. Among this group 18% of men (901) entered initial AS, 42% (2,052) underwent surgery, 21% (1,033) received radiation therapy, 5% (241) were treated with androgen deprivation therapy, 4% (188) elected watchful waiting and 1% (46) received other treatments (eg cryosurgical ablation). Initial therapy was unknown in 9% of these men (425). The 9% of men with missing data on treatment were younger (mean age 63 vs 65 years, $p < 0.001$) and more likely to be black (22% vs 15%, $p < 0.001$), have lower clinical stage tumors (T1 and T2 78% and 20% vs 72% and 26%, respectively, $p = 0.03$) and be from a practice with greater than 10 urologists (51% vs 39%, $p < 0.001$). They were otherwise similar to those with treatment documented in the MUSIC registry.

Overall men entering initial AS were older and more likely to have clinical stage T1 tumors (vs T2 or T3/4), lower median PSA, a lower biopsy Gleason score, fewer positive biopsy cores on biopsy and lower GPC (table 1).

AS selection criteria examined in this analysis comprised specific combinations of routinely available clinical data, including Gleason score, PSA, PSA density, clinical T stage, number of positive biopsy cores and GPC. Table 2 presents a summary of selection criteria for several published guidelines as well as the sensitivity of each set of criteria for identifying men entering AS when applied to the entire cohort. Across all patients the sensitivity of these criteria for identifying patients who actually entered AS ranged from a low of 48.9% (95% CI 46.2–51.6) for University of Toronto criteria to a high of 62.4% (95% CI 58.0–66.7) for JH criteria. University of Toronto criteria identified the greatest absolute number of men entering initial AS (647) and comparatively the NCCN very low risk criteria identified the fewest (290). A total of 14 MUSIC practices had at least 10 patients meeting JH selection criteria. Across these sites the sensitivity of the JH guideline to identify patients who actually received surveillance ranged from 27% to 84% ($p < 0.001$, see figure). Of men missing treatment data only 10% (44) met JH selection criteria.

Table 1. *Characteristics of patients undergoing initial AS vs other treatment strategies*

	Initial AS		Other Treatments		p Value
Mean age/median (range)	65/66	(39–87)	65/65	(38–95)	0.02
No. race (%):					
White	670	(84)	2,589	(82)	
Black	98	(12)	460	(15)	0.1
Other	29	(4)	88	(3)	
No. CCI (%):					
0	595	(68)	2,282	(65)	
1	145	(16)	697	(20)	0.05
2 or Greater	145	(16)	526	(15)	
No. clinical stage (%):					
T1	791	(88)	2,404	(68)	<0.001
T2	104	(12)	1,050	(30)	
T3/4	0	(0)	88	(2)	
Median PSA (ng/ml)	5.3		5.9		<0.001
No. biopsy Gleason score sum (%):					
6 or Less	724	(81)	796	(23)	<0.001
3 + 4	135	(15)	1,375	(39)	
4 + 3	21	(2)	585	(17)	
8–10	10	(1)	732	(21)	
Mean No. pos cores/ median (range)	1.9/1	(0–11)	4.6/4	(1–24)	<0.001
Mean GPC/median (range)	17.3/10	(0–100)	48.1/46	(0–100)	<0.001

For the cohort of men meeting JH selection criteria for initial AS and who were eligible for treatment (ie those not undergoing watchful waiting or receiving primary androgen deprivation therapy) we compared demographics and cancer characteristics between those who actually entered AS and those who received definitive therapy (table 3). Compared with men undergoing initial AS those receiving initial local therapy were younger ($p < 0.001$), had lower PSA ($p < 0.001$) and were more likely to have a positive family history ($p = 0.003$). Treated men also had evidence of greater tumor volume, as indicated by 2 (vs 1) positive cores on biopsy ($p < 0.001$) and higher GPC (median 10% vs 7%, $p = 0.003$). There were no statistically significant differences between these 2 groups for any other variables analyzed, including race and comorbidity.

Among the patients undergoing definitive therapy despite meeting the JH guideline 68% and 32% underwent surgery and radiation, respectively. Those treated with radiation rather than surgery were older (mean age 66 vs 59 years, $p < 0.001$), had more comorbid conditions (CCI 1 and 2 or greater in 33% and 17% vs 11% and 12%, respectively, $p = 0.004$) and were less likely to have a family history of prostate cancer (26% vs 52%, $p = 0.006$).

DISCUSSION

We examined data from a large number of community and academic urology practices to understand the degree to which various published AS selection criteria actually identify men who enter surveillance in real world practice. Not unexpectedly the sensitivity of these guidelines is variable with those containing more stringent criteria generally capturing a greater proportion of all men meeting the criteria who actually entered AS. Conversely

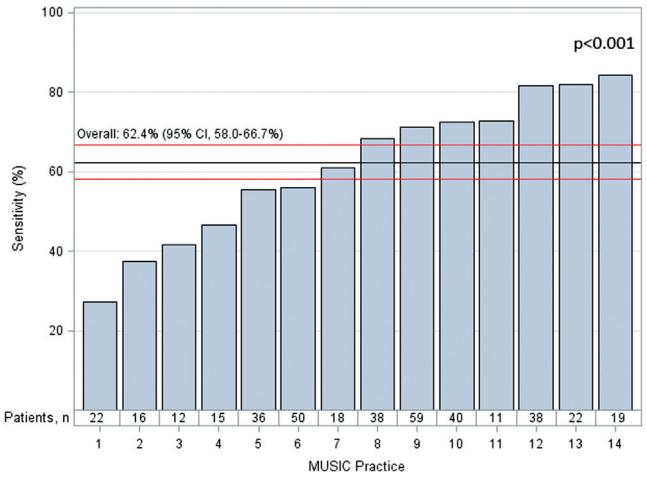
Table 2. *Sensitivity of published AS selection criteria among men in Michigan with newly diagnosed prostate cancer*

Guideline	Selection Criteria*					No. Pts Meeting Selection Criteria	% Sensitivity (95% CI)
	PSA (ng/ml)	PSA Density (ng/ml/ml)	T Stage	No. Pos Cores	% GPC		
JH	—	Less than 0.15	cT1c	2 or Less	50 or Less	486	62.4 (58.0—66.7)
NCCN:							
Very low risk	Less than 10	Less than 0.15	cT1c	2 or Less	50 or Less	466	62.2 (57.8—66.6)
Low risk	Less than 10	—	cT2a or less	—	—	1,271	49.7 (46.9—52.4)
Memorial Sloan Kettering Cancer Center	Less than 10	—	cT2a or Less	3 or Less	Less than 50	984	56.3 (53.2—59.4)
University of California-San Francisco	10 or Less	—	cT2 or Less	33% or Less	50 or Less	1,083	54.3 (51.3—57.3)
University of Toronto	Less than 10	—	—	—	—	1,323	48.9 (46.2—51.6)

* Gleason score 6 or less at each site.

more liberal criteria identified a greater absolute number of men entering AS. The most sensitive guideline overall (JH) still showed significant variation in performance at the practice level. Among the cohort of patients who met JH criteria and were considered eligible for local therapy those who received definitive treatment had a higher tumor volume on biopsy as well as other characteristics favoring treatment (eg they were younger and had a positive family history) that are not explicit components of current AS selection criteria.

In addition, our findings suggest that when urologists counsel men about AS vs definitive treatment, they consider age and family history, which are 2 relevant factors not routinely captured by AS guidelines, as well as small differences in the volume of cancer. Previous work exploring the implications of such differences in tumor volume has been mixed with others reporting that GPC greater than 10% but not 1 vs 2 positive cores at biopsy increased the risk of adverse surgical pathology for patients with Gleason 6 prostate cancer.[17] While our findings certainly seem reassuring with respect to current practice patterns, the clinical significance of these differences in tumor volume is admittedly debatable.



Sensitivity variation of JH guideline to identify men on AS across MUSIC practices with greater than 10 patients meeting selection criteria. Overall sensitivity was 62.4%. Variability among practices was statistically significant ($p < 0.001$).

Our overall findings are consistent with previous investigations showing that AS guidelines vary in the number of men with newly diagnosed prostate cancer who meet such selection criteria.[13] Moreover, the observed practice level variation in the sensitivity of AS guidelines is consistent with our prior reports of AS among patients with low risk cancers[5] as well as prior studies demonstrating variation in primary treatment of localized prostate cancer by practice site and provider.[3,18] While our study identified statistically significant differences in measures of cancer risk between patients who met JH criteria and entered surveillance vs those who received initial local therapy, the clinical significance of these differences is modest at best. As such, it seems likely that variation in AS among men who meet the stringent JH selection criteria may reflect primarily differences in patient preferences or provider perceptions and beliefs about AS that were not measured in this analysis.

The optimal selection criteria for AS remain uncertain. Using more stringent criteria would likely identify patients at lower risk for disease progression but it may also exclude some men who are good candidates for surveillance. Our findings also indicate that urologists in Michigan have not coalesced around a single set of selection criteria for this important initial treatment decision.

Our analysis has several limitations. 1) Our cohort included only patients and practices in

Table 3. *Patients meeting JH criteria who received initial AS vs definitive local therapy*

	Entered AS		Local Therapy		p Value
No. pts	303		132		
Mean age/median (range)	64/65 (41—83)		61/62 (41—77)		<0.001
Median PSA (ng/ml)	5		4.5		<0.001
No. pos cores (%):					
1	226	(75)	76	(58)	<0.001
2	77	(25)	56	(42)	
Median GPC (range)	7	(1—50)	10	(1—50)	0.002
No. pos family history (%)	82	(28)	53	(44)	0.003

Michigan and, therefore, our findings may not be generalizable to a broader population. 2) Treatment data are missing on a small number of men in the cohort. Although such missing data raises concerns about selection bias, even if all 44 men with missing treatment data did not receive AS, this would only decrease the sensitivity of the JH criteria to 57%. Conversely if all of these men entered initial AS, the sensitivity of the JH criteria would increase to 65%. 3) We measured neither patient preferences nor physician beliefs and perceptions about the risks of prostate cancer, treatment side effects and comorbidities. These unmeasured factors undoubtedly have a role in treatment decisions, including the selection of patients for AS. 4) Although our study examined initial entry into AS, the effectiveness of AS ultimately depends on long-term use with select intervention for men with disease progression. Therefore, further study is needed to better define the outcomes of AS in this cohort.

Despite these limitations our findings have important implications for patients and providers. For patients our findings suggest opportunities to expand the use of AS, particularly among men who meet selection criteria and who prioritize preservation of urinary and/or sexual function. For providers these data suggest that even in a single state there appear to be substantial differences in beliefs and perceptions around prostate cancer risk (among other factors) that affect treatment recommendations for patients with low risk tumors. As suggested by others, differing interpretations of the evidence base supporting the safety of AS may explain the disconnection between providers who routinely put patients on AS and those who do not.[19] It is also possible that differences in other unmeasured factors impact treatment decisions, such as financial considerations, practice setting (eg urban vs rural) and/or the presence and strength of any academic affiliation.

CONCLUSIONS

Moving forward, a better understanding of the entire decision making process is needed. This includes characterization of provider perceptions of risk and thresholds for treatment, how cancer risk is communicated with patients, shared decision making between patients and providers, and the degree to which differences in patient preferences drive the variation observed in this analysis. Once available, such data will provide an essential context for understanding the implications of existing variation in the use of AS for men with low risk prostate cancer.

ACKNOWLEDGMENTS

The clinical champions, urologists and data abstractors at each participating MUSIC practice, as well as members of the MUSIC Coordinating Center at the University of Michigan, contributed to this study. Dr. David Share, Tom Leyden, Rozanne Darland and Sarah Lanivich, Value Partnerships, Blue Cross Blue Shield of Michigan, provided assistance.

REFERENCES

1. Daskivich TJ, Chamie K, Kwan L et al: Overtreatment of men with low-risk prostate cancer and significant comorbidity. Cancer 2011; **117:** 2058.

2. Miller DC, Gruber SB, Hollenbeck BK et al: Incidence of initial local therapy among men with lower-risk prostate cancer in the United States. J Natl Cancer Inst 2006; **98:** 1134.

3. Cooperberg MR, Broering JM and Carroll PR: Time trends and local variation in primary treatment of localized prostate cancer. J Clin Oncol 2010; **28:** 1117.

4. Jacobs BL, Zhang Y, Skolarus TA et al: Growth of high-cost intensity-modulated radiotherapy for prostate cancer raises concerns about overuse. Health Aff 2012; **31:** 750.

5. Womble PR, Montie JE, Ye Z et al: Contemporary use of initial active surveillance among men in Michigan with low-risk prostate cancer. Eur Urol 2015; **67:** 44.

6. 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.

7. National Comprehensive Cancer Institute: Prostate Cancer (1.2015). Available at http://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf. Accessed October 2014.

8. Berglund RK, Masterson TA, Vora KC et al: Pathological upgrading and up staging with immediate repeat biopsy in patients eligible for active surveillance. J Urol 2008; **180:** 1964.

9. Welty CJ, Cowan JE, Nguyen H et al: Extended followup and risk factors for disease reclassification in a large active surveillance cohort for localized prostate cancer. J Urol 2015; **193:** 807.

10. Dall'Era MA, Konety BR, Cowan JE et al: Active surveillance for the management of prostate cancer in a contemporary cohort. Cancer 2008; **112:** 2664.

11. 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.

12. Klotz L: Active surveillance for prostate cancer: for whom? J Clin Oncol 2005; **23:** 8165.

13. Dall'Era MA, Albertsen PC, Bangma C et al: Active surveillance for prostate cancer: a systematic review of the literature. Eur Urol 2012; **62:** 976.

14. Womble PR, Dixon MW, Linsell SM et al: Infection related hospitalizations after prostate biopsy in a statewide quality improvement collaborative. J Urol 2014; **191:** 1787.

15. Riedinger CB, Womble PR, Linsell SM et al: Variation in prostate cancer detection rates in a statewide quality improvement collaborative. J Urol 2014; **192:** 373.

16. Merdan S, Womble PR, Miller DC et al: Toward better use of bone scans among men with early-stage prostate cancer. Urology 2014; **84:** 793.

17. Katz MH, Shikanov S, Sun M et al: Gleason 6 prostate cancer in one or two biopsy cores can harbor more aggressive disease. J Endourol 2011; **25:** 699.

18. Hoffman KE, Niu J, Shen Y et al: Physician variation in management of low-risk prostate cancer. JAMA Intern Med 2014; **174:** 1450.

19. Kim SP, Gross CP, Nguyen PL et al: Perceptions of active surveillance and treatment recommendations for low-risk prostate cancer: results from a national survey of radiation oncologists and urologists. Med Care 2014; **52:** 579.