Association Between Radiation Therapy, Surgery, or Observation for Localized Prostate Cancer and Patient-Reported Outcomes After 3 Years

Daniel A. Barocas, MD, MPH; JoAnn Alvarez, MA; Matthew J. Resnick, MD, MPH; Tatsuki Koyama, PhD; Karen E. Hoffman, MD, MHSc, MPH; Mark D. Tyson, MD; Ralph Conwill, BS; Dan McCollum, BS; Matthew R. Cooperberg, MD, MPH; Michael Goodman, MD, MPH; Sheldon Greenfield, MD; Ann S. Hamilton, PhD, MA; Mia Hashibe, PhD, MPH; Sherrie H. Kaplan, PhD, MS, MPH; Lisa E. Paddock, PhD, MPH; Antoinette M. Stroup, PhD; Xiao-Cheng Wu, MD, MPH; David F. Penson, MD, MPH

IMPORTANCE Understanding the adverse effects of contemporary approaches to localized prostate cancer treatment could inform shared decision making.

OBJECTIVE To compare functional outcomes and adverse effects associated with radical prostatectomy, external beam radiation therapy (EBRT), and active surveillance.

DESIGN, SETTING, AND PARTICIPANTS Prospective, population-based, cohort study involving 2550 men (≤80 years) diagnosed in 2011-2012 with clinical stage cT1-2, localized prostate cancer, with prostate-specific antigen levels less than 50 ng/mL, and enrolled within 6 months of diagnosis.

EXPOSURES Treatment with radical prostatectomy, EBRT, or active surveillance was ascertained within 1 year of diagnosis.

MAIN OUTCOMES AND MEASURES Patient-reported function on the 26-item Expanded Prostate Cancer Index Composite (EPIC) 36 months after enrollment. Higher domain scores (range, 0-100) indicate better function. Minimum clinically important difference was defined as 10 to 12 points for sexual function, 6 for urinary incontinence, 5 for urinary irritative symptoms, 5 for bowel function, and 4 for hormonal function.

RESULTS The cohort included 2550 men (mean age, 63.8 years; 74% white, 55% had intermediate- or high-risk disease), of whom 1523 (59.7%) underwent radical prostatectomy, 598 (23.5%) EBRT, and 429 (16.8%) active surveillance. Men in the EBRT group were older (mean age, 68.1 years vs 61.5 years, \( P < .001 \)) and had worse baseline sexual function (mean score, 52.3 vs 65.2, \( P < .001 \)) than men in the radical prostatectomy group. At 3 years, the adjusted mean sexual domain score for radical prostatectomy decreased more than for EBRT (mean difference, −11.9 points; 95% CI, −15.1 to −8.7). The decline in sexual domain scores between EBRT and active surveillance was not clinically significant (−4.3 points; 95% CI, −9.2 to 0.7). Radical prostatectomy was associated with worse urinary incontinence than EBRT (−18.0 points; 95% CI, −20.5 to −15.4) and active surveillance (−12.7 points; 95% CI, −16.0 to −9.3) but was associated with better urinary irritative symptoms than active surveillance (5.2 points; 95% CI, 3.2 to 7.2). No clinically significant differences for bowel or hormone function were noted beyond 12 months. No differences in health-related quality of life or disease-specific survival (3 deaths) were noted (99.7%-100%).

CONCLUSIONS AND RELEVANCE In this cohort of men with localized prostate cancer, radical prostatectomy was associated with a greater decrease in sexual function and urinary incontinence than either EBRT or active surveillance after 3 years and was associated with fewer urinary irritative symptoms than active surveillance; however, no meaningful differences existed in either bowel or hormonal function beyond 12 months or in other domains of health-related quality-of-life measures. These findings may facilitate counseling regarding the comparative harms of contemporary treatments for prostate cancer.

Corrected on May 23, 2017.

© 2017 American Medical Association. All rights reserved.
The optimal management for localized prostate cancer depends on factors including risk of progression; competing risks of mortality, baseline urinary, sexual, and bowel function; and patient preferences. Comparing the effectiveness and harms of radiation therapy, radical prostatectomy, and active surveillance is critical for shared decision making. Yet comparative data have limited generalizability for several reasons, such as focusing on homogenous populations and comparing older treatments instead of contemporary robotic radical prostatectomy and intensity-modulated radiotherapy (IMRT).

In this context, the Comparative Effectiveness Analysis of Surgery and Radiation (CEASAR) study, a prospective, longitudinal, population-based cohort study was developed. In light of the nearly 100% 5-year survival for men with localized prostate cancer, patient-reported disease-specific functional outcomes were selected as the primary short- and intermediate-term outcome measures. This study assessed patient-reported functional outcomes and health-related quality of life at 3 years after treatment.

Methods

The parent study accrued men diagnosed with localized prostate cancer (2011-2012) from 5 Surveillance, Epidemiology, and End Results (SEER) registries (Atlanta [Georgia], Los Angeles [California], Louisiana, New Jersey, and Utah), and the Cancer of the Prostate Strategic Urologic Research Endeavor registry. Details of the protocol have been published. Eligibility criteria were being younger than 80 years, having a prostate-specific antigen (PSA) level of less than 50 ng/mL, clinical stage T1 to T2 cancer, no nodal involvement or metastases on clinical evaluation; and being enrolled within 6 months of diagnosis (Table).

Patient-reported outcomes were collected via mail survey at enrollment and 6, 12, and 36 months after enrollment. If patients did not respond to 2 mailings, trained abstractor called the patient to complete the survey. A medical chart review, including clinical and treatment information, was obtained at 12 months. SEER registry data were linked to the data set. This study includes follow-up through August 2015. Institutional review board approval was obtained from each site and from Vanderbilt University Medical Center. Patients provided written informed consent.

Outcomes

The primary outcome measures were 36-month domain scores on the 26-item Expanded Prostate Cancer Index Composite (EPIC-26), a validated instrument for measuring disease-specific function. Domain scores range from 0 to 100, with higher score representing better function. The minimal clinically important difference (MCID), representing the magnitude of change that is clinically meaningful to patients, has been estimated for each domain using standard techniques. The distribution-based approach estimated MCID as one-third to one-half of a standard deviation, and the anchoring approach identified the magnitude of change on each domain that resulted in a change in satisfaction with treatment. Both techniques yielded similar MCIDs and were consistent with the a priori definition of MCID used in the power calculation of the original grant application for this study, which was one-half of a standard deviation. The sexual function domain focuses on the quality and frequency of erections (MCID, 10-12 points). The urinary incontinence (MCID, 6 points) and urinary irritative symptom (MCID, 5 points) domains ask questions about frequency; amount of urinary leakage; and dysuria, hematuria, and urinary frequency. The bowel function domain (MCID, 4 points) focuses on bowel frequency, urgency, bleeding, and pain. The hormonal domain (MCID, 4 points) assesses symptoms such as hot flashes, gynecomastia, low energy, and weight change. The baseline survey asked about pretreatment function. Previous studies have investigated the issue of recall bias for the EPIC instrument, including a study in this cohort, and adjusted differences in domain scores between those who complete the survey before treatment and those who complete it afterward range from 1.0 to 3.7 points, well below the MCID for each domain.

Individual items from the EPIC-26 were selected a priori as secondary outcomes based on clinical relevance by content experts and patients on the study team.

Treatments were also compared with respect to health-related quality of life, using selected domains from the commonly used Medical Outcomes Study Short Form 36 (SF-36): physical functioning, emotional well-being, and energy and fatigue. Domain scores are scaled from 0 to 100, with higher scores indicating better function. The MCIDs for these domains have been estimated for patients with localized prostate cancer as 7, 6, and 9 points, respectively.

Exposure

The main exposure was initial treatment, defined according to the following hierarchy of sources: medical chart abstraction, patient report, and SEER registry. A participant was categorized as undergoing active surveillance if this strategy was documented in the absence of treatment or if no treatment was administered within 1 year of diagnosis.
Table. Demographics and Baseline Characteristics

<table>
<thead>
<tr>
<th>Demographics</th>
<th>No. (%) of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Radical Prostatectomy (n = 1523)</td>
</tr>
<tr>
<td>Age, mean (95% CI), y</td>
<td>61.5 (61.1-61.8)</td>
</tr>
<tr>
<td>Race/ethnicity, No.</td>
<td>1511</td>
</tr>
<tr>
<td>White</td>
<td>1130 (75)</td>
</tr>
<tr>
<td>Black</td>
<td>187 (12)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>125 (8)</td>
</tr>
<tr>
<td>Asian</td>
<td>46 (3)</td>
</tr>
<tr>
<td>Other</td>
<td>23 (2)</td>
</tr>
<tr>
<td>Education, No.</td>
<td>1411</td>
</tr>
<tr>
<td>High school</td>
<td>130 (9)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>302 (21)</td>
</tr>
<tr>
<td>Some college</td>
<td>315 (22)</td>
</tr>
<tr>
<td>College graduate</td>
<td>345 (24)</td>
</tr>
<tr>
<td>Graduate or professional school</td>
<td>349 (24)</td>
</tr>
<tr>
<td>Marital status, No.</td>
<td>1138</td>
</tr>
<tr>
<td>Married</td>
<td>1192 (83)</td>
</tr>
<tr>
<td>Comorbidity score, No.</td>
<td>1448</td>
</tr>
<tr>
<td>0-2</td>
<td>481 (33)</td>
</tr>
<tr>
<td>3-4</td>
<td>624 (43)</td>
</tr>
<tr>
<td>5</td>
<td>343 (24)</td>
</tr>
<tr>
<td>D'Amico prostate cancer risk, No.</td>
<td>1521</td>
</tr>
<tr>
<td>Low risk</td>
<td>635 (42)</td>
</tr>
<tr>
<td>Intermediate risk</td>
<td>635 (42)</td>
</tr>
<tr>
<td>High risk</td>
<td>251 (17)</td>
</tr>
<tr>
<td>Prostate specific antigen, ng/mL, No.</td>
<td>1523</td>
</tr>
<tr>
<td>0-4</td>
<td>334 (22)</td>
</tr>
<tr>
<td>4.1-10</td>
<td>1018 (67)</td>
</tr>
<tr>
<td>10.1-20</td>
<td>133 (9)</td>
</tr>
<tr>
<td>&gt;20</td>
<td>38 (2)</td>
</tr>
<tr>
<td>Clinical stage, No.</td>
<td>1520</td>
</tr>
<tr>
<td>T1c</td>
<td>1140 (75)</td>
</tr>
<tr>
<td>T2</td>
<td>380 (25)</td>
</tr>
<tr>
<td>Biopsy Gleason score, No.</td>
<td>1519</td>
</tr>
<tr>
<td>3 + 3</td>
<td>744 (49)</td>
</tr>
<tr>
<td>3 + 4</td>
<td>458 (30)</td>
</tr>
<tr>
<td>4 + 3</td>
<td>170 (11)</td>
</tr>
<tr>
<td>8-10</td>
<td>147 (10)</td>
</tr>
<tr>
<td>Any hormone therapy in the first year, No.</td>
<td>1509</td>
</tr>
<tr>
<td>Yes</td>
<td>75 (5)</td>
</tr>
<tr>
<td>Survey scores</td>
<td></td>
</tr>
<tr>
<td>Expanded Prostate Cancer Index Composite^c</td>
<td>1447</td>
</tr>
<tr>
<td>Sexual domain, No. of patients</td>
<td>65.2 (63.5-66.9)</td>
</tr>
<tr>
<td>Urinary incontinence, No. of patients</td>
<td>1467</td>
</tr>
<tr>
<td>Mean (95% CI)</td>
<td>86.7 (85.5-87.8)</td>
</tr>
<tr>
<td>Urinary irritation, No. of patients</td>
<td>1463</td>
</tr>
<tr>
<td>Mean (95% CI)</td>
<td>83.2 (82.3-84.1)</td>
</tr>
<tr>
<td>Bowel function, No. of patients</td>
<td>1492</td>
</tr>
<tr>
<td>Mean (95% CI)</td>
<td>94.0 (93.3-94.6)</td>
</tr>
<tr>
<td>Hormonal function, No. of patients</td>
<td>1467</td>
</tr>
<tr>
<td>Mean (95% CI)</td>
<td>89.8 (89.1-90.5)</td>
</tr>
</tbody>
</table>

(continued)
Distinguishing between watchful waiting, active surveillance, and treatment delay was not possible. We categorized these patients as being actively surveilled recognizing that it was a heterogeneous group. For analysis, time 0 was the date patients underwent either radical prostatectomy or EBRT; whereas, the date of diagnosis was time 0 for patients who were being actively surveilled.

**Statistical Analysis**

Baseline characteristics were compared across treatments using Kruskal-Wallis tests for continuous variables and \( \chi^2 \) tests for categorical variables.

To describe typical trajectories of function over time, longitudinal regression models were fit to predict EPIC domain scores as a function of treatment, time since treatment, and their interaction. For each domain, a single model was fit incorporating domain scores from all time points. We used generalized estimating equations (GEE) with an independent weight matrix because of the correlation between observations on the same patients. Modeling time using regression splines allowed for a flexible relationship between function and time. Variability in the interval between treatment and survey completion allowed for estimation of domain scores between rounds of data collection, and beyond 36 months.

Recognizing that outcomes (and patients' priorities) may differ by baseline function, we repeated these models, stratifying by baseline domain scores (excellent and less than excellent). Because excellent function has not been defined in the literature based on EPIC domain scores, a cutoff baseline score was selected for each domain that approximated the highest quartile of domain scores, an approach that has been used in prior publications on patient-reported outcomes after prostate cancer treatment.20

<table>
<thead>
<tr>
<th>Table. Demographics and Baseline Characteristics (continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medical Outcomes Study</strong></td>
</tr>
<tr>
<td>Short Form 36 score[^d]</td>
</tr>
<tr>
<td>Physical function, No. of patients</td>
</tr>
<tr>
<td>Mean (95% CI)</td>
</tr>
<tr>
<td>Emotional well-being, No. of patients</td>
</tr>
<tr>
<td>Energy or fatigue, No. of patients</td>
</tr>
<tr>
<td>Modified Social Support Scale score, No. of patients[^a]</td>
</tr>
<tr>
<td>Mean (95% CI)</td>
</tr>
<tr>
<td>Center for Epidemiologic Studies score[^f]</td>
</tr>
<tr>
<td>Depression scale, No. of patients</td>
</tr>
<tr>
<td>Mean (95% CI),</td>
</tr>
<tr>
<td>Medical decision-making style, No. of patients[^g]</td>
</tr>
<tr>
<td>Mean (95% CI)</td>
</tr>
<tr>
<td>Accrual site, No. of patients</td>
</tr>
<tr>
<td>Louisiana</td>
</tr>
<tr>
<td>Utah</td>
</tr>
<tr>
<td>Atlanta</td>
</tr>
<tr>
<td>Los Angeles County, California</td>
</tr>
<tr>
<td>New Jersey</td>
</tr>
<tr>
<td>Cancer of the Prostate</td>
</tr>
<tr>
<td>Strategic Urologic Research Endeavor</td>
</tr>
<tr>
<td>122 (8) 26 (4)</td>
</tr>
</tbody>
</table>

[^a]: Based on the Total Illness Burden Index, scores range from 0 to 23, with higher score indicating greater severity and number of comorbid illnesses.

[^b]: D’Amico risk classification system predicts risk of recurrence after treatment. Low-risk disease is defined as a clinical stage T2a or less, Gleason score 6 (3 + 3) or less, and a prostate-specific antigen (PSA) less than 10 ng/mL; high-risk disease, T2c or higher, Gleason score 8 (3 + 5, 4 + 4, 5 + 3) or greater, or a PSA value greater than 20 ng/mL; intermediate risk not defined by low- or high-risk definition.

[^c]: Scores range from 0 to 100, with higher scores representing better function.

[^d]: The Medical Outcomes Short-Form Health Survey 36 (SF-36) has 8 domains. The physical function domain score is a weighted sum of 10 items; emotional well-being, 5 items; energy or fatigue score, 4 items. Each domain score is directly transformed to a scale of 0 to 100 with increasing scores indicating better function or less disability.

[^e]: Five questions were selected to create a modified domain score. Responses were directly transformed to a scale of 0 to 100 with increasing scores indicating greater support.

[^f]: Derived using the 10-item Center for Epidemiologic Studies Depression Scale. Scores were scaled to 100, with higher scores indicating more severe depressive symptoms.

[^g]: Seven items were scored on a scale from 0 to 100, with higher scores indicating increased patient choice, control, and responsibility.
To measure the association between treatment choice and domain score over time, a similar set of models was fit that adjusted for age, race/ethnicity, comorbidity, prostate cancer risk stratum, physical function, social support, depression, medical decision-making style, site, and baseline EPIC domain score. This multivariable modeling approach was designed to minimize bias associated with known differences in baseline characteristics that are associated with functional outcomes (ie, confounding). Multiple imputation was used for missing covariates (see eMethods in the Supplement). Because androgen deprivation therapy is a standard component of radiation therapy for high-risk disease and an option in intermediate-risk disease, androgen deprivation therapy was not controlled for in the models. Instead, exploratory models were fit for sexual and hormonal function with 5 treatment groups: nerve-sparing radical prostatectomy, non–nerve-sparing radical prostatectomy, EBRT without androgen deprivation therapy, EBRT with androgen deprivation therapy, and active surveillance. Unadjusted and adjusted longitudinal regression models using GEE were fit for responses to individual EPIC items and for the 3 SF-36 domains, using the same covariates as above. In the SF-36 regression models, the baseline SF-36 domain score was added as an independent variable.

Probability of overall and disease-specific survival was estimated by treatment using the Kaplan-Meier technique with log-rank tests.

Differences in domain scores between treatments were statistically significant if the 2-tailed P value was < .05. Domain scores were interpreted as clinically meaningful if the differences were as large as the MCID. R version 3.2.2 was used for all analyses.

Results

The parent study accrued 3709 patients, of whom 440 were excluded for failing to meet basic inclusion criteria. An additional 519 were excluded from the current study for receiving a treatment other than radical prostatectomy, EBRT, or active surveillance, leaving 2750 patients for consideration (eFigure 1 in the Supplement). The analytic cohort contained 2550 men (93%) who completed a baseline survey and at least 1 survey thereafter. Approximately 93% of surveys were completed on paper, while 7% were completed by telephone; 98% of surveys were conducted in English and 2% in Spanish; 54% of baseline surveys were collected prior to initial treatment. Survey response rates were 89% at 6 months, 86% at 12 months, and 78% at 36 months (eFigure 1 and eTable 1 in the Supplement).

Among men in the analytic cohort, 1523 (59.7%) underwent radical prostatectomy, 598 (23.5%) EBRT, and 429 (16.8%) active surveillance. Baseline characteristics are shown in the Table. Briefly, 26% of the cohort was non-white. Patients treated with EBRT were older, had higher comorbidity burden, and had higher-risk disease features than did patients who were treated with radical prostatectomy. Seventy-seven percent of active surveillance patients had low-risk disease.

Of the 1302 men (71%) who underwent radical prostatectomy and had complete reporting of nerve-sparing status, 859 (79%) had bilateral nerve-sparing, and of the 1032 (85%) who had complete reporting of the surgical approach, 1002 (77%) had robotic surgery. Of the 593 patients (99%) treated with EBRT who had complete reporting of utilization of androgen deprivation therapy, 265 (45%) received androgen deprivation therapy within the first year diagnosis of treatment; 378 patients (81%) of the 467 with complete records underwent IMRT. By the 3-year survey, 24.2% of active surveillance patients had undergone treatment, and 90.2% of the remainder had their PSA checked within the past 12 months.

For the stratified analyses, 26.4% of patients had excellent baseline domain scores for sexual function (≥ 90 points), 26.1% for urinary irritative symptoms (100 points), 61.7% for bowel function (100 points), and 39.1% for hormonal function (100 points).

Sexual Function

Patients undergoing radical prostatectomy had higher baseline sexual domain scores than men undergoing EBRT and had scores comparable with those on active surveillance (eTable 2 in the Supplement). Radical prostatectomy and EBRT were associated with declines in sexual function scores, but the decline was greater for patients who underwent radical prostatectomy, resulting in similar average unadjusted domain score for radical prostatectomy and EBRT at 3 years (Figure 1A, C, and D). The difference in functional decline between radical prostatectomy and EBRT was greater for the 26.4% of men with excellent baseline function, while the 73.6% of men with lower baseline function had poor sexual function outcomes regardless of whether they underwent radical prostatectomy or EBRT. Active surveillance was associated with preservation of function, with mild decline over time.

When controlling for baseline domain scores and other covariates (eTable 2; Figure 1B), men who underwent radical prostatectomy had a larger decline in sexual domain score than did those who underwent EBRT (adjusted mean domain score difference at 3 years, −11.9 points; 95% CI, −15.1 to −8.7) or active surveillance (−16.2; 95% CI, −20.6 to −11.7), relative to the MCID of 10 to 12 points. The adjusted mean domain score after EBRT was significantly worse than it was for active surveillance at 12 months (−10.5; 95% CI, −14.0 to −6.9), but the magnitude of difference at 3 years was no longer significant (−4.3; 95% CI, −9.2 to 0.7). Treatment, baseline domain score, and time since treatment were the only variables for which the magnitude of association with 3-year domain scores exceeded the MCID.

On exploratory analysis with a 5-tier treatment variable (nerve-sparing radical prostatectomy, non–nerve-sparing radical prostatectomy, EBRT alone, EBRT plus androgen deprivation therapy, and active surveillance), the mean difference between EBRT alone and active surveillance was not statistically significant (−3.0 points, P = .27), and the mean difference...
Figure 1. Association Between Initial Treatment of Prostate Cancer and Sexual Function Outcomes

A) All men

B) Adjusted difference in sexual function domain score at 3 y

C) Men with excellent baseline domain score (≥90 points)

D) Men with lower baseline domain score (<90 points)

E) Erection sufficient for intercourse

F) Sexual function bother (reporting no, very small, or small problem)

See the Methods section and the Table footnotes for definitions measures, scoring ranges, and minimum clinically important differences. Baseline is defined as the date of initiation of treatment (in men undergoing radical prostatectomy or radiation) or date of diagnosis for those who were actively surveilled. The number of patients represent those who completed surveys after enrollment. Shaded areas indicate 95% CIs.

B. Data in the forest plots are estimated by multivariable regression models that controlled for baseline domain score, age, race, comorbidity, prostate cancer risk group, physical function, social support, depression, decision-making style, and accrual site. Positive values represent better outcomes for the nonreferent group. eTable 8 in the Supplement contains unadjusted domain scores and number of patients for each subgroup by treatment. See the Table footnotes for definition of disease risk.
between radical prostatectomy and EBRT plus androgen deprivation therapy was attenuated (−8.2 points; 95% CI, −13.2 to −3.2) lower than the MCID (eFigure 2 in the Supplement).

More men who underwent radical prostatectomy were bothered by sexual dysfunction 3 years after diagnosis (44% vs 35% for EBRT and 28% for active surveillance, P < .001 on multivariable analysis; Figure 1F; eTable 2 in the Supplement). Erection insufficient for intercourse was common at 3 years (70% for radical prostatectomy, 71% for EBRT, and 51% for active surveillance on raw percentages, Figure 1E), but when controlling for baseline sexual function and other factors, the odds were significantly higher for radical prostatectomy than for active surveillance (odds ratio [OR], 3.4; 95% CI, 2.5 to 4.6) and for radical prostatectomy than EBRT (OR, 2.1; 95% CI, 1.5 to 2.9). Among men who had sufficient erections at baseline, erection sufficient for intercourse at 3 years was reported in 43% (95% CI, 40% to 47%) of men who had undergone radical prostatectomy; 53% (95% CI, 45% to 60%) EBRT; and 75% (95% CI, 68%–80%), active surveillance, in raw percentages. An exploratory multivariable model, using 5 treatment groups, yielded similar results (eTable 3 in the Supplement).

Urinary Incontinence
Baseline urinary incontinence domain scores were similar across groups (eTable 4 in the Supplement). However, radical prostatectomy was associated with a significant decline in urinary incontinence score after treatment, particularly in the 60.3% of men with perfect urinary incontinence domain scores at baseline (Figure 2A, C, and D). There was no significant change in urinary incontinence score for men who had EBRT or active surveillance, regardless of baseline score.

Despite some improvement in incontinence domain scores 12 months after radical prostatectomy, adjusted mean incontinence scores were still significantly worse for radical prostatectomy than for active surveillance (−12.7 points, 95% CI, −16.0 to −9.3) and EBRT (−18.0 points, 95% CI, −20.5 to −15.4) at 3 years, differences greater than the MCID of 5 points (Figure 2B; eTable 4 in the Supplement). By contrast, urinary incontinence was not significantly different between EBRT and active surveillance. Treatment, baseline domain score, and time since treatment were the only variables for which the magnitude of association with the 3-year domain score exceeded the MCID.

Reports of moderate or big problems with urinary leakage were more common after radical prostatectomy vs active surveillance (14% vs 6%; OR, 2.9; 95% CI, 1.8–4.7) and radical prostatectomy than EBRT (14% vs 5%; OR, 4.5; 95% CI, 2.7–7.3; Figure 2E; and eTable 4 in the Supplement). Urinary function bother scores were not significantly different for radical prostatectomy vs active surveillance and EBRT vs active surveillance at 3 years but were higher for radical prostatectomy vs EBRT (12% vs 10%; OR, 1.7; 1.1–2.5; Figure 2F; and eTable 4 in the Supplement).

Urinary Irritative Symptoms
Baseline scores were similar across groups (eTable 4 in the Supplement). Scores improved for radical prostatectomy, particularly for the 73.9% of men whose baseline score was less than 100 (Figure 3A, C, and D). Those undergoing EBRT or active surveillance experienced little or no change in irritative urinary symptoms.

Adjusted urinary irritative function scores were slightly better for men undergoing radical prostatectomy than being actively surveilled at 1 year (4.5 points; 95% CI, 3.0–6.0) and 3 years (5.2 points, 95% CI, 3.2–7.2), at the threshold of clinical significance (Figure 3B; eTable 4 in the Supplement). Other comparisons across treatments, while statistically significant, were lower than the MCID of 5 (Figure 3B; eTable 4 in the Supplement). Besides treatment with radical prostatectomy, the only other factors for which the magnitude of association with 3-year domain score exceeded the MCID were baseline domain score and time since treatment.

Reports of moderate or big problems with burning with urination were uncommon (2% in each group; Figure 3E; eTable 4 in the Supplement). Reports of moderate or big problem with frequent urination were lower for radical prostatectomy than for active surveillance (13% vs 18%; OR, 0.6; 95% CI, 0.4–0.8) and for EBRT vs active surveillance (15% vs 18%, OR, 0.6; 95%; 0.4–0.8) at 3 years, but not significantly different between radical prostatectomy and EBRT (Figure 3F; eTable 4 in the Supplement).

Bowel Function
Decline in bowel domain score was not common (Figure 4A, C, and D; eTable 5 in the Supplement). Six months after treatment, the mean domain scores were higher in men who underwent radical prostatectomy than who underwent EBRT (4.6 points, 95% CI, 3.2–6.1) and lower for EBRT vs active surveillance (−5.8 points; 95% CI, −10.3 to −1.2 points). However, by 12 months these differences were near the MCID of 4 and by 36 months, they were smaller. Unadjusted and adjusted results were similar (Figure 4B). No other independent variables had a magnitude of association with 3-year domain score that met the threshold for clinical significance.

The frequency of moderate or big problem with bowel bother, bloody stools, or bowel urgency was 1% to 8% across all treatments at 3 years (Figure 4E–F; eTable 5 in the Supplement). Nevertheless, the odds of bowel urgency at 3 years were lower for radical prostatectomy than EBRT (3% vs 7%, OR, 0.3; 95% CI, 0.2–0.6) and radical prostatectomy vs active surveillance (3% vs 5%, OR, 0.5; 95%, 0.3–0.9).

Hormone Function
The mean hormone domain scores were worse for EBRT than for active surveillance and radical prostatectomy at 6 months (radical prostatectomy vs EBRT, 5.0 points; 95% CI, 3.3 to 6.6 points; EBRT vs active surveillance, −6.5 points; 95% CI, −11.1 to −1.9), but these differences no longer significant at 3 years on unadjusted or adjusted analyses (Figure 5; eTable 6 in the Supplement). No other independent variables had a magnitude of association with 3-year domain score that reached the MCID.

In the exploratory models that separated EBRT into with and without androgen deprivation therapy, the only group with decrements in hormone function was the EBRT plus...
Figure 2. Association Between Initial Treatment of Prostate Cancer and Urinary Incontinence Outcomes

A. All men

B. Adjusted difference in urinary incontinence domain score at 3 y

C. Men with excellent baseline domain score (100 points)

D. Men with lower baseline domain score (<100 points)

E. Urinary leakage (reporting no, very small, or small problem)

F. Urinary function bother (reporting no, very small, or small problem)

No. of patients
Active surveillance | 409 | 408 | 374 | 339
External beam radiation | 575 | 576 | 532 | 467
Radical prostatectomy | 1467 | 1434 | 1368 | 1277

No. of patients
Active surveillance | 240 | 228 | 208 | 192
External beam radiation | 337 | 325 | 300 | 282
Radical prostatectomy | 901 | 849 | 830 | 790

No. of patients
Active surveillance | 415 | 410 | 385 | 347
External beam radiation | 582 | 579 | 551 | 474
Radical prostatectomy | 1492 | 1442 | 1411 | 1293

See the Methods section for explanation of the measures, scoring ranges, and minimum clinically important difference definitions. Baseline is defined as the date of initiation of treatment (in men undergoing radical prostatectomy or radiation) or date of diagnosis for those who were actively surveilled. The number of patients indicates those who completed surveys at the specified time following enrollment. Shaded areas indicate 95% CIs. See Figure 1 legend for an explanation of the forest plot.
Figure 3. Association Between Initial Treatment of Prostate Cancer and Urinary Irritative Outcomes

A. All men

No. of patients
- Active surveillance: 409, 403, 379, 340
- External beam radiation: 574, 571, 547, 466
- Radical prostatectomy: 1463, 1419, 1397, 1283

B. Adjusted difference in urinary irritative symptoms domain score at 3 y

- Radical prostatectomy vs active surveillance [referent]
- External beam radiation vs active surveillance [referent]
- Baseline domain score (75th vs 25th percentile [referent])
- Age at diagnosis (70 vs 60 y [referent])
- D’Amico intermediate risk vs low risk [referent]
- D’Amico high risk vs low risk [referent]

C. Men with excellent baseline domain score (100 points)

No. of patients
- Active surveillance: 101, 93, 88, 78
- External beam radiation: 121, 115, 113, 101
- Radical prostatectomy: 417, 381, 380, 360

D. Men with lower baseline domain score (<100 points)

No. of patients
- Active surveillance: 409, 403, 379, 340
- External beam radiation: 574, 571, 547, 466
- Radical prostatectomy: 1463, 1419, 1397, 1283

E. Burning on urination (reporting no, very small, or small problem)

No. of patients
- Active surveillance: 413, 409, 385, 347
- External beam radiation: 582, 576, 554, 477
- Radical prostatectomy: 1488, 1439, 1410, 1298

F. Frequent urination (reporting no, very small, or small problem)

No. of patients
- Active surveillance: 414, 409, 382, 347
- External beam radiation: 585, 577, 555, 479
- Radical prostatectomy: 1491, 1440, 1413, 1299

See the Methods section for explanation of the measures, scoring ranges, and minimum clinically important difference definitions. Baseline is defined as the date of initiation of treatment (in men undergoing radical prostatectomy or radiation) or date of diagnosis for those who were actively surveilled. The number of patients indicate those who completed surveys at the specified time following enrollment. Shaded areas indicate 95% CIs. See Figure 1 legend for an explanation of the forest plot.
Figure 4. Association Between Initial Treatment of Prostate Cancer and Bowel Function Outcomes

A. All men

No. of patients
- Active surveillance: 415, 411, 385, 349
- External beam radiation: 585, 576, 557, 476
- Radical prostatectomy: 1492, 1439, 1415, 1297

B. Adjusted difference in bowel function domain score at 3 y

- Radical prostatectomy vs active surveillance [referent]
- External beam radiation vs active surveillance [referent]
- Baseline domain score (75th vs 25th percentile [referent])
- Age at diagnosis (70 vs 60 y [referent])
- D'Amico intermediate risk vs low risk [referent]
- D'Amico high risk vs low risk [referent]

C. Men with excellent baseline domain score (100 points)

No. of patients
- Active surveillance: 259, 246, 234, 213
- External beam radiation: 331, 319, 309, 273
- Radical prostatectomy: 948, 896, 892, 821

D. Men with lower baseline domain score (<100 points)

No. of patients
- Active surveillance: 156, 152, 138, 125
- External beam radiation: 254, 245, 235, 193
- Radical prostatectomy: 544, 514, 498, 452

E. Bowel urgency (reporting no, very small, or small problem)

No. of patients
- Active surveillance: 416, 411, 385, 349
- External beam radiation: 586, 578, 558, 478
- Radical prostatectomy: 1492, 1441, 1415, 1300

F. Bowel function bother (reporting no, very small, or small problem)

No. of patients
- Active surveillance: 415, 410, 385, 349
- External beam radiation: 586, 576, 557, 477
- Radical prostatectomy: 1477, 1439, 1415, 1301

See the Methods section for explanation of the measures, scoring ranges, and minimum clinically important difference definitions. Baseline is defined as the date of initiation of treatment (in men undergoing radical prostatectomy or radiation) or date of diagnosis for those who were actively surveilled. The number of patients indicate those who completed surveys at the specified time following enrollment. Shaded areas indicate 95% CIs. See Figure 1 legend for an explanation of the forest plot.
androgen deprivation therapy group, and these associations were limited to the first year (eFigure 2 in the Supplement).

**Health-related Quality of Life**
Baseline physical functioning and energy or fatigue scores on the SF-36 were lower for men undergoing EBRT than radical prostatectomy or active surveillance (eTable 7 in the Supplement). None of the treatment groups experienced a clinically significant decline in physical functioning, emotional well-being, or energy or fatigue scores (Figure 6). On multivariable analysis, associations between treatment and 3-year SF-36 quality-of-life domain scores were below the threshold for clinical significance, as were associations baseline EPIC sexual and urinary incontinence domain scores and 3-year SF-36 domain scores.

**Survival**
Median follow-up time among censored patients was 40 months (interquartile range [IQR], 38-45, months). There were 78 deaths, including 3 prostate cancer deaths. On Kaplan-Meier analysis, estimated 3-year disease-specific survival was not significantly different across groups (99.9% for radical prostatectomy, 99.7% for EBRT, and 100% for active surveillance). Unadjusted 3-year overall survival was higher for radical prostatectomy (99%, 95% CI, 98%-99%) than for other groups (EBRT, 96%; 95% CI, 94%-98%; active surveillance, 97%; 95% CI, 95-99; P < .001), commensurate with the younger age and lower comorbidity of men undergoing radical prostatectomy (eTable 9 in the Supplement).

**Discussion**
In this study of men with localized prostate cancer, radical prostatectomy was associated with clinically significant declines in sexual function compared with EBRT and active...
Figure 6. Association Between Initial Treatment of Prostate Cancer and Overall Quality-of-Life Outcomes

A  Physical functioning

Quality of Life Score, Unadjusted Mean (95% CI)

0 20 40 60 80 100

Baseline 6 12 18 24 30 36 Time Since Treatment Start, mo

Active surveillance  
External beam radiation  
Radical prostatectomy

No. of patients
Active surveillance  405 409 387 347
External beam radiation  577 578 558 480
Radical prostatectomy  1477 1442 1419 1398

B  Emotional well-being

Quality of Life Score, Unadjusted Mean (95% CI)

0 20 40 60 80 100

Baseline 6 12 18 24 30 36 Time Since Treatment Start, mo

Active surveillance  413 410 386 346
External beam radiation  586 577 556 476
Radical prostatectomy  1492 1442 1412 1297

C  Energy or fatigue

Quality of Life Score, Unadjusted Mean (95% CI)

0 20 40 60 80 100

Baseline 6 12 18 24 30 36 Time Since Treatment Start, mo

Active surveillance  414 410 386 346
External beam radiation  586 577 556 476
Radical prostatectomy  1493 1441 1412 1398

See the Methods section for explanation of the Medical Outcomes Study measures, scoring ranges, and minimum clinically important difference definitions. Baseline is defined as the date of initiation of treatment (in men undergoing radical prostatectomy or radiation) or date of diagnosis for those who were actively surveilled. Solid lines indicate unadjusted means; shaded areas, 95% CIs. Shaded areas indicate 95% CIs. See Figure 1 legend for an explanation of the forest plot.
surveillance, particularly among men with excellent function at baseline. Urinary incontinence scores also declined significantly after surgery compared with EBRT and active surveillance, with 14% of patients treated with radical prostatectomy reporting a moderate or big problem with urinary leakage at 3 years compared with 5% with EBRT and 6% with active surveillance. Radical prostatectomy was associated with better irritative voiding symptoms than active surveillance, with a difference that met the threshold for clinical significance. Mean scores in bowel and hormonal domains were significantly worse for EBRT vs radical prostatectomy and active surveillance at 6 months, but the differences were below threshold for clinical significance by 3 years. Treatment, baseline domain scores, and time since treatment were the independent variables with clinically significant associations with 3-year domain scores. None of the treatment groups experienced clinically significant declines in health-related quality-of-life domain scores. This information may facilitate patient counseling regarding the expected harms of contemporary treatments and their possible effect on quality of life.

Prior studies have quantified the harms of prostate cancer treatment. However, randomized trials studying localized prostate cancer have been difficult to execute, and those that have been completed focused on outdated treatments; enrolled too few minority patients; lacked a range of disease severity; failed to collect baseline functional assessments; or included a preponderance of elderly, infirm, and low-risk patients, for whom treatment is questionable.° The ProtecT trial, for example, included 99% white patients and nearly 80% of patients with a Gleason score of 6 (low-risk). In ProtecT, 87% of surgical patients underwent open radical prostatectomy (vs 77% who underwent robotic surgery in this study) and patients undergoing EBRT had 3-dimensional conformal radiation therapy plus androgen deprivation therapy (compared with 81% receiving IMRT, with 45% receiving concurrent androgen deprivation therapy in this study). Thus, the ProtecT study findings may be difficult to apply to a racially diverse population with a range of disease risk strata, managed with contemporary treatments.

Case series that have evaluated functional outcomes are not generalizable because they reported on outcomes at centers of excellence; lacked the variables necessary to adjust for confounding; lacked an active surveillance group as a comparator; or had other sources of bias.° Despite these caveats, functional outcomes in this study are similar to previously published multi-institutional prospective cohort studies and to the ProtecT trial. Nevertheless, comparisons between the CEASAR cohort and similar historical cohorts have shown slightly smaller declines in erectile function domain scores at 6 and 12 months with robotic radical prostatectomy than with open radical prostatectomy, and slightly better bowel domain scores at 6 months for IMRT than for older 3-dimensional conformal radiation therapy.° These data suggest that contemporary treatments have similar associations with functional outcomes but perhaps slightly less in magnitude.

This study may have implications for decision making for patients with localized prostate cancer. First, it demonstrates the frequency and severity of adverse effects of contemporary treatments and the likelihood of preserved global quality of life regardless of treatment, thus providing a basis for shared decision making. Second, in contrast to previously published studies, this study may be more generalizable, since the cohort is racially diverse, population based, and includes a range of disease severity.° Third, this study may inform future research on personalized risk assessment, tools to facilitate shared decision making, and other patient-centered outcomes.

Limitations

This study has several limitations. There may be disagreement about the definition of MCID, which may also differ from one patient to the next. Although some outcomes favored one treatment over another, the results do not indicate what value patients place on particular domains. Furthermore, there are other important outcomes to consider in localized prostate cancer, including long-term functional outcomes and oncologic end points, anxiety, satisfaction, and financial toxicity. The number and severity of adverse outcomes presenting beyond 3 years may differ by treatment, and 3 years is inadequate to estimate oncologic outcomes. Data on patients who had other treatments, such as brachytherapy and ablation, were not included because there were not enough patients who received these treatments to generate sufficient statistical power for reliable comparisons. Aggregated data and average function scores may fail to capture the severity of adverse effects for individuals and do not yield personalized risk estimates. The analysis did not adjust for the quality of care or experience of the treating clinician or institution, which may influence outcomes. Thus, the findings of this study represent a subset of the information needed to guide decision making. A substantial proportion of patients answered the baseline survey after initiating treatment, raising the possibility of recall bias, although in prior studies the magnitude of recall bias was small for the EPIC survey.° This study used an observational cohort, rather than an experimental design, so there may be unmeasured sources of confounding.

Conclusions

In this cohort of men with localized prostate cancer, radical prostatectomy was associated with a greater decrease in sexual function and urinary incontinence than either EBRT or active surveillance after 3 years and was associated with fewer urinary irritative symptoms than active surveillance; however, no meaningful differences existed in either bowel or hormonal function beyond 12 months or in other domains of health-related quality of life measures. These findings may facilitate counseling regarding the comparative harms of contemporary treatments for prostate cancer.
Outcomes of Radiation, Surgery, or Observation for Localized Prostate Cancer

ARTICLE INFORMATION

Corrected: This article was corrected online May 23, 2017, for incorrect information in the Results section of the article.

Author Affiliations: Department of Urologic Surgery, Vanderbilt University Medical Center, Nashville, Tennessee (Barocas, Resnick, Tyson, Penson); Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee (Alvarez, Koyama); Department of Radiation Oncology, University of Texas MD Anderson Cancer Center, Houston (Hoffman); Prostate Cancer Patient Advocate, Vanderbilt Ingram Cancer Center, Nashville, Tennessee (Conwill, McCollum); Department of Urology, University of California, San Francisco Medical Center, San Francisco (Cooperberg); Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia (Goodman); Center for Health Policy Research and Department of Medicine, University of California, Irvine (Greenfield); Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (Hamilton); Department of Family and Preventive Medicine, University of Utah, Salt Lake City (Hashibe); Health Policy Research Institute, University of California, Irvine (Kaplan); Rutgers Cancer Institute of New Jersey, Rutgers University, New Brunswick (Paddock, Stroup); School of Public Health, Louisiana State University Health Sciences Center, New Orleans (Wu); Tennessee Valley Veterans Administration Health System, Nashville, Tennessee (Penson).

Author Contributions: Drs Barocas and Penson had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Barocas, Koyama, Conwill, Tyson, Greenfield, Kaplan, Paddock, Tyson, Hoffman, Conwill, Song, Penson.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Barocas, Alvarez, Hoffman, Tyson, Conwill, Penson.

Critical revision of the manuscript for important intellectual content: Alvarez, Resnick, Koyama, Conwill, McCollum, Cooperberg, Goodman, Greenfield, Hamilton, Hashibe, Kaplan, Paddock, Stroup, Wu, Penson.

Statistical analysis: Alvarez, Resnick, Koyama, Greenfield, Kaplan, Penson.

Obtained funding: Stroup, Penson.

Administrative, technical, or material support: Resnick, Tyson, Conwill, Goodman, Hamilton, Hashibe, Paddock, Stroup, Wu, Penson.

Supervision: Barocas, Hamilton, Paddock, Stroup, Penson.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Cooperberg reported receiving grants from Myriad Genetics, Genomic Health, and Genome Dx and personal fees from Dendreon and Astellas for serving as a cochair of a registry project and from Bayer and Janssen for serving as an ad board participant. Dr Paddock reports receiving a subcontract from Vanderbilt University to Rutgers Cancers Institute of New Jersey. Dr Resnick reports receiving grants from the American Cancer Society and from the Urology Care Foundation, and personal fees from MDx Health and Janssen Pharmaceuticals. No other disclosures were reported.

Funding/Support: Funding for the study was provided by grants IRO1HS019356 and IRO1HS022640 from the Agency for Healthcare Research and Quality; U10TR000011 to the Vanderbilt Institute of Clinical and Translational Research from the National Center for Advancing Translational Sciences; and ST32CA106183 from the National Institute of Health and the National Cancer Institute (Dr Tyson). Research reported in this article was partially funded through a Patient-Centered Outcomes Research Institute (PCORI) award CE12-11-4667.

Role of the Funder/Sponsor: Each funding organization provided financial support through grants, but none was involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Disclaimer: The statements in this article are solely the responsibility of the authors and do not necessarily represent the views of the PCORI or its board of governors or methodology committee.

REFERENCES


© 2017 American Medical Association. All rights reserved.


