Comparing Treatments for Clinically Localized Prostate Cancer: Review and Evidence Visualization

Prepared by

Meera Viswanathan, PhD[1]

Sheila V. Patel, PhD[1]

Shivani Reddy, MD, MSc[1]

Shauna Durbin, MPH[2]

Edwin Reid, MS, MFA[3]

Scott McDonald, BA[3]

Jeanne-Marie Guise, MD, MPH[3]

Joshua Fenton, MD, MPH[2]

[1] RTI International, Research Triangle Park, North Carolina

[2] University of California-Davis, Sacramento, California

[3] Oregon Health & Sciences University, Portland, Oregon

December 2020

Suggested Citation

Viswanathan M, Patel S, Reddy S, et al. *Comparing Treatments for Clinically Localized Prostate Cancer: Review and Evidence Visualization*. Patient-Centered Outcomes Research Institute; December 2020. Prepared by RTI under Contract No. IDIQ-TO#11-RTI-EVIDENCEMAPAMPTESP.

Table of Contents

Background

Prostate cancer is the most common cancer in men,[1,2] and the vast majority of men who have it are diagnosed with early-stage, clinically localized prostate cancer confined to the prostate gland (stage T1-T2).[2,3] Screening with prostate-specific antigen (PSA) often identifies early-stage disease, is performed commonly in developed countries, and currently has a "C" recommendation from the US Preventive Services Task Force (USPSTF), meaning that clinicians may order PSA screening for men who have a preference to pursue screening.[4] Early-stage prostate cancer can follow an indolent course and many men may never become symptomatic.[5] However, once prostate cancer has spread beyond the prostate, men can experience painful bone metastases, urinary obstruction, weight loss, and other constitutional symptoms associated with cancer. Five-year survival rates drop from 100% for localized disease to 30% for distant disease relative to cancer-free men.[2] The decision to treat early-stage prostate cancers involves considering, on the one hand, the possibility of experiencing common side effects of treatment for a cancer that may never cause symptoms or reduce survival and, on the other, the risk of developing symptomatic disease or death. Although age, African American race, and family history increase risk for more aggressive prostate cancer,[6,7] it is unclear which men have early prostate cancer that will progress to symptomatic metastatic disease and which men have indolent cancers that will not ultimately be the cause of death.

Once diagnosed with early-stage prostate cancer, patients and their providers face the practical decision of which treatment option to pursue and how to balance the immediate harms of active therapy (such as surgery or radiation) with the possibility of delayed harms (disease progression and possible increased mortality) of conservative observation-based options. The patient's first choice is whether to pursue or defer active treatment. Options for deferring active treatment (ie, conservative therapy) include watchful waiting or active surveillance. With watchful waiting, further testing and treatment are performed only when symptoms arise, whereas active surveillance maintains a curative intent with the patient

undergoing regular evaluation using serial PSA measurements and digital rectal exams, and usually biopsy to uncover evidence of asymptomatic disease progression. If there are signs of progression, patients on active surveillance are offered active treatments.

Active treatment options include radical prostatectomy or radiation therapy (with external beam radiation therapy or brachytherapy, in which radioactive seeds are placed in the prostate). Both radical prostatectomy and radiation therapy can result in harms that can potentially affect quality of life (QOL), including erectile dysfunction, urinary incontinence, and bowel dysfunction.[8] Active surveillance may be associated with repeat biopsy-related harms and psychological harms, and for a portion of patients who will experience cancer progression, ultimate treatment with radical prostatectomy or radiation therapy. Although the recent Prostate Testing for Cancer and Treatment Trial (ProtecT), a randomized controlled trial (RCT) comparing active surveillance, radical prostatectomy, and external beam radiation therapy, found no significant decrease in all-cause or prostate cancer mortality for radical prostatectomy or external beam radiation therapy when compared with active surveillance, the point estimates were below 1 (eg, consistent with a lower mortality risk) and there was increased risk of disease progression and metastatic disease in men randomized to active surveillance.[9,10] Additional therapies considered for early-stage prostate cancer include combining radiation therapy with androgen deprivation therapy and focal therapies such as cryotherapy, high-intensity focused ultrasound, and photodynamic therapy.[11-13] For androgen deprivation therapy, harms of erectile dysfunction, hot flashes, and osteoporosis must be considered,[14] and although focal therapy may theoretically spare surrounding organs and reduce harms, much of the evidence is limited to cohort or case-control studies.[12]

In the absence of clear clinical effectiveness of one treatment over another and the risk for harms that significantly affect QOL, treatment of prostate cancer falls under the category of a preference-sensitive decision, amenable to shared decision making. In a joint guideline for the treatment of clinically localized cancer, the American Urological Association, the American Society for Radiation Oncology, and Society of Urologic Oncology give a strong recommendation for incorporating shared decision making into management decisions.[15]

Uncertainty in the evidence on comparative effectiveness, coupled with the important role of patient preference, points to the importance of tools that communicate the weight of the evidence. In this project, the research team systematically identified and evaluated the evidence and laid out the results in the evidence visualizations meant to serve as an evidence entry point for patients with newly diagnosed early-stage prostate cancer and a tool to support patient-provider discussion. This evidence visualization lays out the amount and strength of the evidence when comparing treatment options for 5 key outcomes: all-cause mortality, prostate cancer–specific mortality, urinary function, bowel function, and sexual function. This pilot project explores the use of visual tools to convey complex information; as a result, the visualization focuses on evidence patterns across all patients, including a range of clinical subgroups like age, race, and Gleason score. However, individual patient and provider decision making requires careful consideration of such risk factors. While the visualizations do not depict evidence for clinical subgroups, they do allow users to link to specific studies in the database underlying the evidence map, serving as a resource for further exploration. This document describes the methods, the results, and the resultant visualizations.

Methods

Selection and Engagement of Technical Expert Panel

To ensure that this project met the needs of PCORI and its stakeholders, RTI assembled a technical expert panel. We spoke with 4 clinical experts (3 urologists and 1 radiation oncologist) with ongoing research and clinical practice guideline expertise. We sought their input on the audience for the evidence visualization; framework for treatment decisions; perceived uncertainties in clinical practice; population, intervention, comparator, outcome, timing, and study design (PICOTS) parameters for the review (appendix 1); and potential filters for the evidence visualization. Table 1 lists their names and affiliations at the time of engagement on this project.

Table 1. Clinical Expert Panel

Name	Affiliation
Peter R. Carroll, MD, MPH	University of California, San Francisco
Philipp Dahm, MD, MHSc	University of Minnesota
Nancy Mendenhall, MD	University of Florida
David Penson, MD, MPH	Vanderbilt University

Peer Review

Four peer reviewers (2 urologists and 2 radiation oncologists) reviewed this report. We revised the report in response to their comments.

Literature Search

Our initial review of systematic reviews revealed 2 up-to-date systematic reviews that are consistent with PCORI's substantive questions; one was supported by the USPSTF and authored by Fenton et al[8,16] and the other was commissioned by the Agency for Healthcare Research and Quality (AHRQ) Effective Healthcare Program and authored by Sun et al.[17] The Fenton[8,16] review focused on comparisons of treatments for early-stage prostate cancer with watchful waiting, active surveillance, or other conservative management approaches for early-stage prostate cancer and excluded studies comparing 2 active treatments. The Sun review[17] included comparisons of active treatments in addition to comparisons of treatments against conservative management.

We created search strings to replicate the literature yield and to be consistent with the search criteria in Fenton et al (2017)[8] and Sun et al (2014).[17] Fenton et al last updated their searches in July 2017; Sun et al did so in March 2014. To ensure that the review and evidence visualization cover all the relevant evidence, we first ran the searches from January 1, 2014, to November 27, 2018. We updated the searches on October 9, 2019. Our search dates therefore include a small overlap with Sun et al (January through March 2014) and a larger overlap with Fenton et al (January 2014 through July 2017). Appendix 1 lists our search strings for PubMed,

Cochrane, and gray literature sources, specifically, HSRPproj, clinicaltrials.gov, and AHRQ and PCORI projects.

Inclusion and Exclusion Criteria

Although our inclusion and exclusion criteria were generally consistent with the 2 reviews that served as our base, one exception was that we excluded single-arm studies on focal therapy. The focus of the evidence visualization is on comparative effectiveness; single-arm focal therapy studies cannot contribute to that evidence base. Table 2 lists our inclusion and exclusion criteria. Key criteria include (1) a focus on clinically localized prostate cancer (T1 to T2 disease only), (2) exclusion of primary androgen deprivation therapy, and (3) exclusion of within-class comparisons (for example, we did not include studies comparing different types of radiation therapy). Our key informant interviews and internal expertise suggested greater evidence and consensus for active treatment of more invasive disease and avoiding androgen deprivation therapy as a primary treatment.[18,19] As patients are among consumers of this evidence visualization, we compared only broad treatment categories; we considered more granular decisions like surgical techniques and radiation modality or dose to be outside the scope of shared decision making.

Table 2. Inclusion and Exclusion Criteria

Criterion	Inclusion	Exclusion
Population	• Men with clinically localized prostate cancer (T1 or T2 disease) • Studies with mixed populations that include ≤10% T3 disease • Studies that do not report T-category but describe the population as having clinically localized cancer	• Men with advanced prostate cancer (T3 disease or higher, positive lymph nodes or distant metastases), refractory or recurrent prostate cancer, or small cell prostate carcinoma • Studies with mixed populations that have >10% with T3 disease and no stratification of results

Criterion	Inclusion	Exclusion
		• Studies that do not report T-stage and the population is described as having high-risk prostate cancer without specifying if clinically localized

Interventions	• Surgery
	  ○ Radical prostatectomy (RP) (robotic, laparoscopic, open)
	• Radiation therapy (RT)
	  ○ External beam radiation therapy (EBRT)
	  ○ Low-dose rate brachytherapy (BT)
	  ○ High-dose rate BT
	  ○ Stereotactic body radiation therapy
	  ○ Proton therapy
	• Focal therapy
	  ○ Cryosurgery
	  ○ High-intensity focused ultrasonography
	  ○ Other focal therapies: photodynamic therapy, magnetic resonance imaging–guided transurethral ultrasound ablation, focal irreversible electroporation
	• Combination therapy
	  ○ RT + hormonal therapy (ADT)
	  ○ EBRT + BT
		• Treatments for advanced prostate cancer (bilateral orchiectomy, chemotherapy, cancer vaccines)
		• Primary androgen deprivation therapy (ADT)
		• RP combined with neo-ADT

Comparator	• Conservative management (active surveillance and/or watchful waiting and/or unspecified) • Watchful waiting or no treatment	• No comparison for benefits or harms • Head-to-head comparisons of different types of RT (including RT vs RT + ADT)

Criterion	Inclusion	Exclusion
	• Active comparators listed above for active surveillance, surgery, RT, combination therapies, or focal therapy	• Head-to-head comparisons of different types of surgery
Outcome	Benefits: • All-cause mortality • Prostate cancer–specific mortality • Metastatic disease (defined as M1) Harms: • Incidence of physical harms of treatment, as measured by validated tools • Erectile dysfunction • Urinary incontinence • Bowel dysfunction • Procedure complications • Biopsy complications (infection, hematuria, pain) • Surgical complications (perioperative mortality, bleeding, thromboembolic disease, rectal lesions, urinary fistulae) • Post-procedure hospitalization • Medication adverse effects • Hot flashes • Gynecomastia • Hepatitis • Osteoporosis • Cardiac risk • Cognitive function • Disease-specific quality of life • Psychological harms, measured by validated tools • Depression • Anxiety • Procedure burden • Decision regret	• Regional disease progression (N1) • Biochemical progression • Hormone levels • Health care costs • Outcomes other than those specified

Criterion	Inclusion	Exclusion
Timing	Published in 2014 or later	Published before 2014
Study design	• Systematic reviews (SRs) of individual studies to identify randomized and nonrandomized controlled trials for benefits[a] • SRs with quality ratings of individual studies to identify randomized and nonrandomized controlled trials and observational studies for harms • Randomized and nonrandomized controlled trials and observational cohorts for benefits and harms • Case-control studies for harms	• Case studies • Case-control studies for benefits • Case series/single-arm studies for benefits or harms of active surveillance, surgery, RT, focal therapy, or combination therapies
Other	English language	Non–English language

[a] We expanded this criterion from SRs with quality ratings to all SRs on further review.

Abstract and Full-Article Screening

For studies identified through the database search, we dually and independently reviewed titles and abstracts against the inclusion and exclusion criteria. Articles flagged for inclusion by either reviewer then moved to full-text review. Each full-text article was then dually and independently reviewed. Consensus or third-party adjudication helped resolve all conflicts.

Risk-of-bias Assessment

To ensure consistent bias assessment across studies retrieved for this review update, we applied the risk-of-bias assessment criteria used in the review by Fenton et al to all other eligible studies identified through other systematic reviews and primary searches.[8] Appendix 2 lists the criteria for RCTs and cohorts. Ratings for studies included by Fenton et al were carried forward for our update. For the sake of efficiency, we single-reviewed the risk-of-bias ratings for studies marked as medium or high quality and dual-reviewed the studies marked as low

quality in the review by Sun et al.[17] Studies that were excluded from Sun et al that met our inclusion criteria were dual-reviewed. Additional studies identified via our update searches were dual-reviewed. All conflicts were resolved by consensus or third-party adjudication. We ultimately excluded studies that we marked as poor quality.

Data Extraction and Analysis

One team member systematically extracted data into Excel sheets. A second reviewer checked for omissions and inaccuracies. Differences or conflicts were resolved by consensus. For each study, we collected first author; year of publication; PubMed identifier (PMID); registry or trial name; related references; study design; country; latest timepoint and nature of the follow-up as described (planned, median, etc); intervention(s); comparator(s); number of subjects by treatment arm; baseline age and measure of age (mean, median, etc); percentage white, black, and other; and baseline risk (including mean or median PSA values; percentage with T1 and T2 scores; percentage with Gleason scores less than 7 vs 7 and greater; and percentage with low, intermediate, and high clinical risk assessment, as defined in Table 3).

Table 3. Risk Classification for Localized Prostate Cancer[20,21]

Risk category[a]	PSA	Gleason Grade Group[b]	Clinical stage[c]
Low risk	<10	2	T1-T2a
Intermediate risk	10 to <20	2-3	T2b-c
High risk	≥20	4	T3a

Abbreviation: PSA, prostate-specific antigen.

[a] For the low-risk category, all criteria in the row must be met. For overall intermediate and high-risk categories, at least one criterion must be met.

[b] Gleason Grade Group categorization endorsed by the World Health Organization and US and Canadian Academy of Pathology.
Grade 1 = Gleason 6
Grade 2 = Gleason 3+4
Grade 3 = Gleason 4+3
Grade 4 = Gleason 4+4
Grade 5 = Gleason 4+5[22]

[c] Clinical stage determined by palpation of lesions on digital rectal exam defined as follows:
T1 = not palpable
T2a = unilateral, <50% lobe involvement
T2b = unilateral, >50% lobe involvement
T2c = bilateral lobe involvement
T3 = extension through prostate capsule (not relevant for this topic brief)

For each eligible outcome in each article, we recorded the first author, year, PMID, registry/trial name, study design, country, number of subjects analyzed for the specific comparison, outcome category, specific outcome, population subgroup, latest timepoint, nature of the follow-up as described (planned, median, etc), and other timepoints reported. For each arm of the comparison, we recorded the number of subjects analyzed, outcome, type of outcome measure (mean, percentage, etc), and crude and adjusted measures of effect along with confidence intervals.

Strength of Evidence Assessment

We evaluated the strength of evidence based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group guidance and guidance established for the AHRQ's Evidence-based Practice Center Program. This approach incorporates 5 key domains: body of evidence limitations, consistency, directness, precision of the evidence, and reporting bias.[23-31]

- Body of evidence limitations account for inherent limitations of study design and the risk of bias of individual studies. For example, we rated bodies of evidence comprising unadjusted results from cohort studies alone as having high study limitations because of their failure to account for confounding. We rated bodies of evidence comprising results from trials as having low or moderate study limitations, depending on the underlying risk of bias of the included studies.
- Consistency judgments for the body of evidence consider consistency across individual studies in the direction and magnitude of the effect.
- Directness judgments for the body of evidence consider the applicability of the population, intervention, comparator, outcome, and so on, in individual studies, to the outcome being graded. Based on our selection criteria and the reported outcomes, we rated all outcomes as direct. As a result, we do not list directness in the strength of evidence tables.
- Precision judgments for the body of evidence evaluate the degree of certainty in the effect estimates. These judgments require looking beyond statistical significance, even when studies are consistent and of high quality and outcomes are direct and clinically relevant. The judgments emphasize the adequacy of the sample size to rule out spurious associations and results that are not clinically relevant.
- Reporting bias considers the likelihood that reported outcomes or analyses may have been selectively published.

- Optional domains that may also influence judgments about strength of evidence include dose-response association, plausible confounding that would decrease the observed effect, and strength of association (magnitude of effect).

These judgments on individual domains are then summarized in a judgment regarding the strength of the overall body of evidence. We assign 1 of 4 grades:

- **High**: This grade indicates high confidence that the evidence reflects the true effect. Further research is *very unlikely* to change our confidence in the estimate of effect.
- **Moderate:** This grade indicates moderate confidence that the evidence reflects the true effect. Further research *may* change our confidence in the estimate of effect and may change the estimate.
- **Low**: This grade indicates low confidence that the evidence reflects the true effect. Further research is *likely* to change our confidence in the estimate of effect and is likely to change the estimate.
- **Insufficient**: The grade indicates that the evidence does not permit a conclusion.

Additionally, we indicate when no evidence is available for a given outcome. In the strength of evidence grades, we generally graded specific comparisons, rather than combining across different types of conservative management comparisons. For example, we graded comparisons of radical prostatectomy vs active surveillance separately from radical prostatectomy vs watchful waiting or unspecified conservative management. Similarly, we graded the evidence for radical prostatectomy vs radiation therapy plus androgen deprivation therapy separately from radical prostatectomy vs radiation therapy alone. The exception was high- and low-dose brachytherapy. For studies presenting different doses of brachytherapy alone, we did not grade separately by brachytherapy dose.

Visualization Construction

Website developers and user interface (UI/UX) designers constructed a variety of visual representations of the data using both proprietary Tableau and open-source D3 software. (D3 uses JavaScript, CSS, and HTML to render scalable vector graphics in a browser window.) These representations were revised iteratively through input from the PCORI and RTI project teams to make visualizations usable by and understandable to the PCORI web audience. Given the nature of these data, and to provide the broadest possible visual display, we developed 3 visualizations. Evidence Visualization 1 gives an overview of all studies in our review and

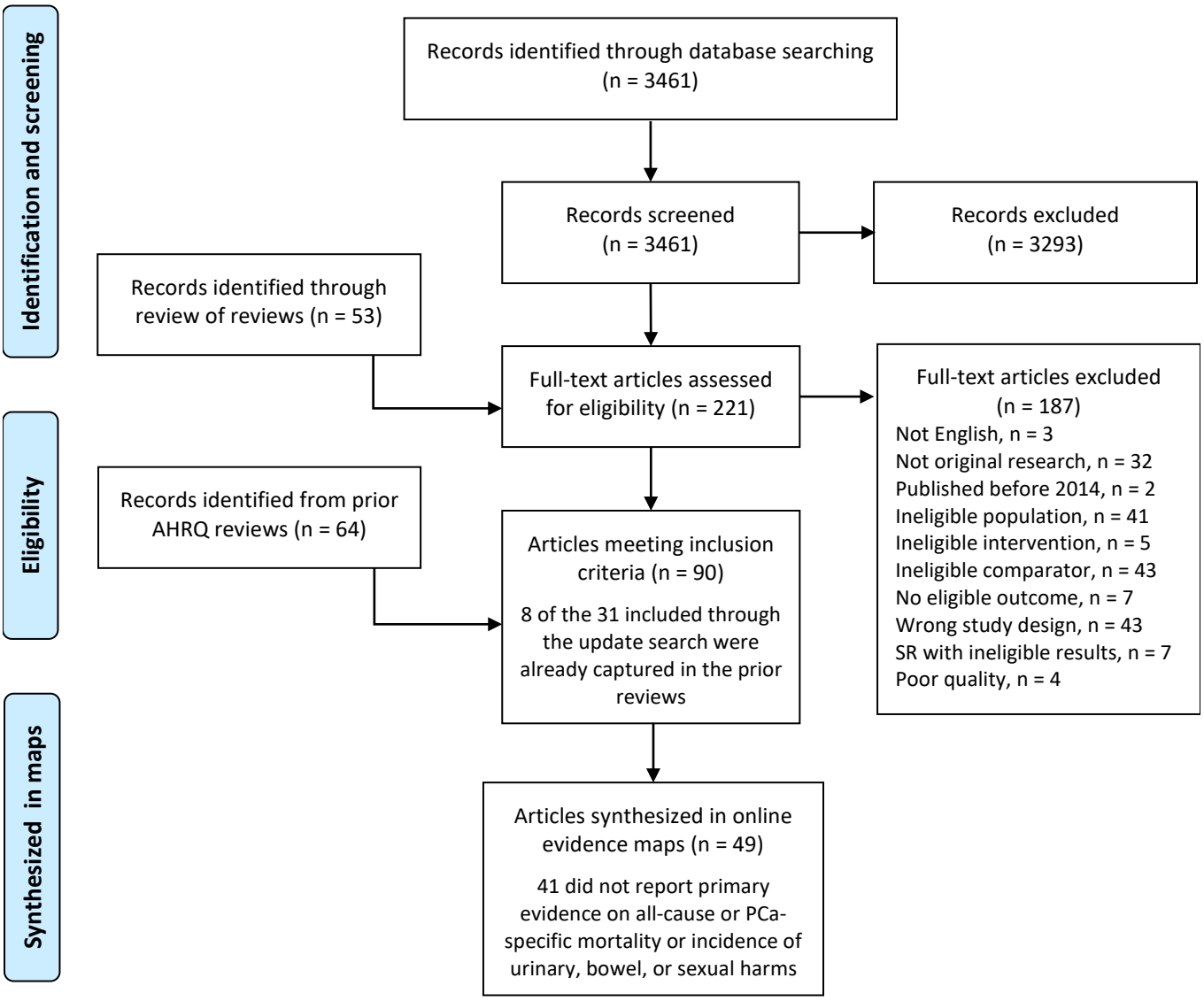
consequently provides data on 90 publications. The Heat Map and Evidence Visualization 2 provide a more detailed look at our strength of evidence assessments. Because early visualizations that included a larger number of outcomes were difficult to follow, the final visualizations are limited to data on 5 outcomes deemed important to patients and clinicians, based on input from the clinicians on the team. These visualizations are based on 49 publications with data on at least one of these outcomes.

Results

Review Yield

Our searches yielded 3461 potentially relevant articles (see article flow in Figure 1) and an additional 53 citations from reviewing citations from relevant systematic reviews. Abstract screening resulted in the exclusion of 3293 articles from all evidence visualizations. Of the 221 full-text articles evaluated for eligibility, 26 met inclusion criteria and added to the base of 64 eligible articles already identified from prior reviews (Fenton et al and Sun et al).[8,17] The total number of included articles was 90. The 3 most common exclusions were due to (1) articles not having an eligible comparator, (2) articles having the wrong study design, and (3) articles having an ineligible population. These 90 articles are depicted in Evidence Visualization 1 and the Heat Map. We then rated the strength of evidence for 5 critical outcomes (all-cause mortality, prostate cancer–specific mortality, urinary function, bowel function, and sexual function). This evidence base comprises 49 articles (comprising 40 studies) that are depicted in Evidence Visualization 2.

Figure 1. Article Flow for Update of AHRQ Systematic Reviews on Treatments for Clinically Localized Prostate Cancer



Abbreviations: AHRQ, Agency for Healthcare Research and Quality; PCa, prostate cancer; SR, systematic review.

Evidence Visualizations

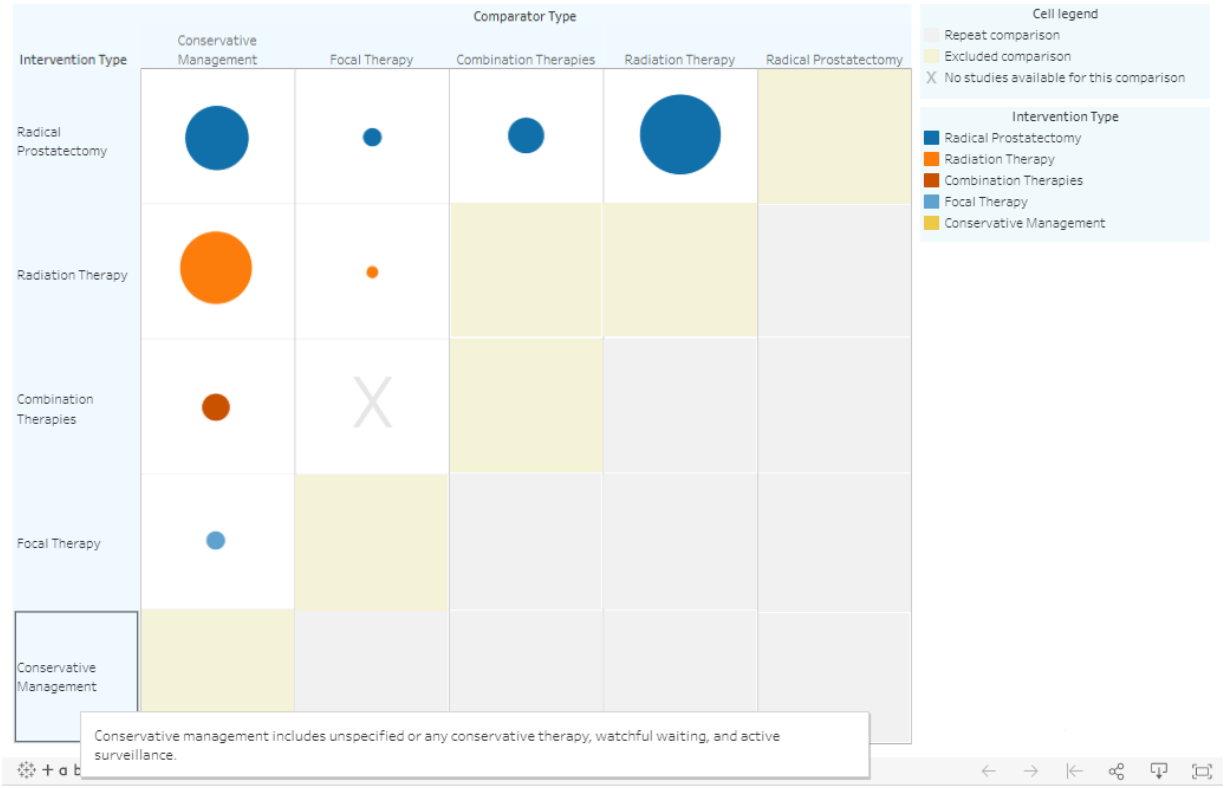
Evidence Visualization 1: Volume of Evidence Available by Treatment Comparison

For Evidence Visualization 1 (see a screenshot in Figure 2), interventions are categorized in a matrix by overall treatment and comparator types. Key visual elements are listed below.

- The **color of the bubble** indicates the intervention type.
- The **size of each bubble** corresponds to the number of studies investigating that treatment and comparison type; this information is also available in a tooltip when hovering over a bubble.
- The **shading of each cell in the matrix** indicates the availability of evidence or eligibility of the comparison treatment.
- **Tooltips for the first column** of the matrix reveal more details on interventions included within each category.

The visualization as a whole depicts the greatest volume of evidence in comparisons of radical prostatectomy vs radiation therapy, radical prostatectomy vs conservative management (watchful waiting, active surveillance, or unspecified), and radiation therapy vs conservative management. Radiation vs combination therapy (radiation plus androgen deprivation therapy) was considered a within-group comparison (variants of radiation therapy) and excluded as a result.

Figure 2. Evidence Visualization 1: Overall Number of Publications for Treatment and Comparator Types (displaying the tooltip for conservative management)

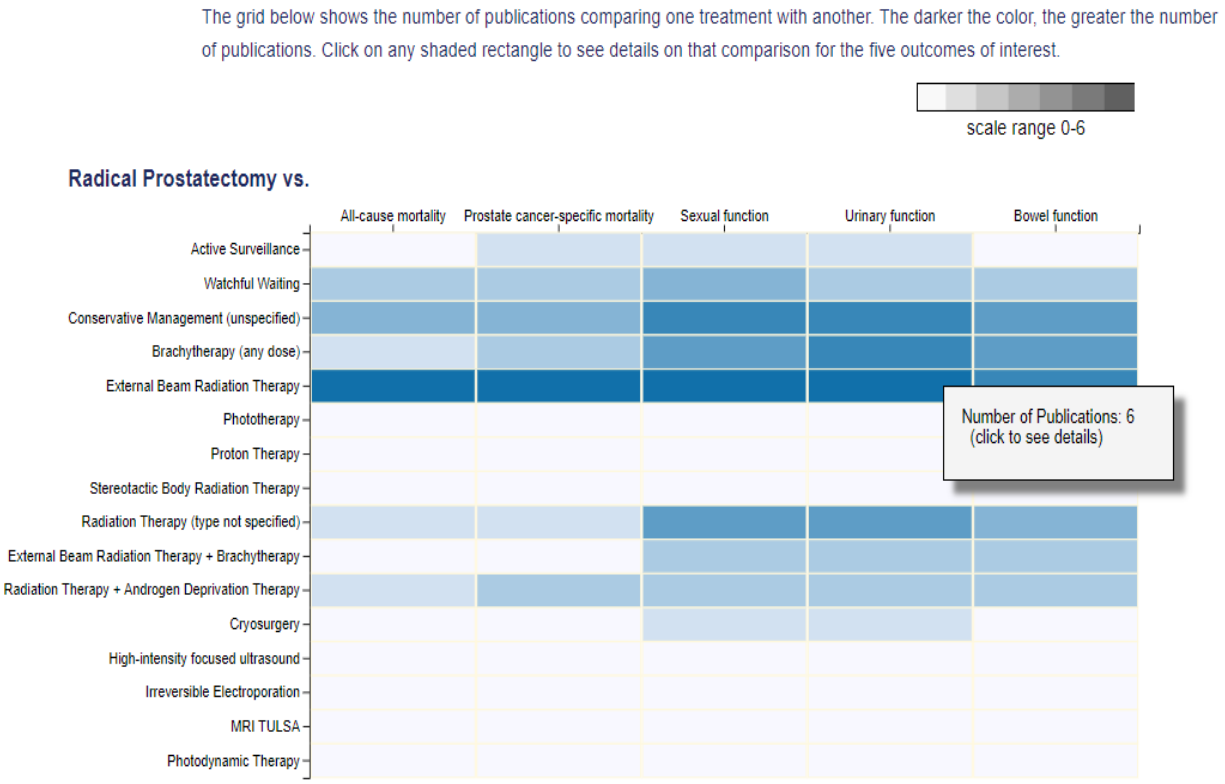


Clicking on any bubble other than radiation therapy vs conservative management leads to Evidence Visualization 2, depicting details on each of the 5 graded outcomes where users can download data on these outcomes; additionally, they can download data on these outcomes in subgroups defined by age, race, and clinical risk. Users, if interested, can download data on other eligible outcomes from this page as well. Because radiation therapy includes many subtypes of interventions, clicking on the radiation therapy vs conservative management bubble leads to another bubble chart showing these subtypes (unspecified, external beam radiation therapy, brachytherapy, and proton therapy). Clicking on a bubble in this secondary bubble chart leads to Evidence Visualization 2.

Heat Map: Volume of Evidence Available by Treatment Comparison

The Heat Map provides an alternative visual of the volume of evidence; **the color of the cell** depicts the number of publications comparing one treatment with another for each of the 5 graded outcomes (see a screenshot in Figure 3). The darker the color, the greater the number of publications. This map shows the specific intervention and comparator in granular detail, revealing the greatest volume of evidence is for comparisons of radical prostatectomy and external beam radiation therapy and for comparisons of unspecified radiation therapy with unspecified conservative management. Clicking on any individual cell leads to Evidence Visualization 2. This map also highlights gaps in the evidence base, including treatments prespecified in the inclusion criteria for which we found no evidence (photon therapy, proton therapy, stereotactic body radiation therapy, high-intensity focused ultrasound, irreversible electroporation, magnetic resonance imaging–guided transurethral ultrasound ablation [MRI-TULSA]).

Figure 3. Heat Map for Radical Prostatectomy and Comparators[a]



[a] This heat map depicts the 5 outcomes of particular interest to patients and clinicians and displays the tooltip over External Beam Radiation Therapy and Urinary function. Other visualizations appear below this map on the webpage.

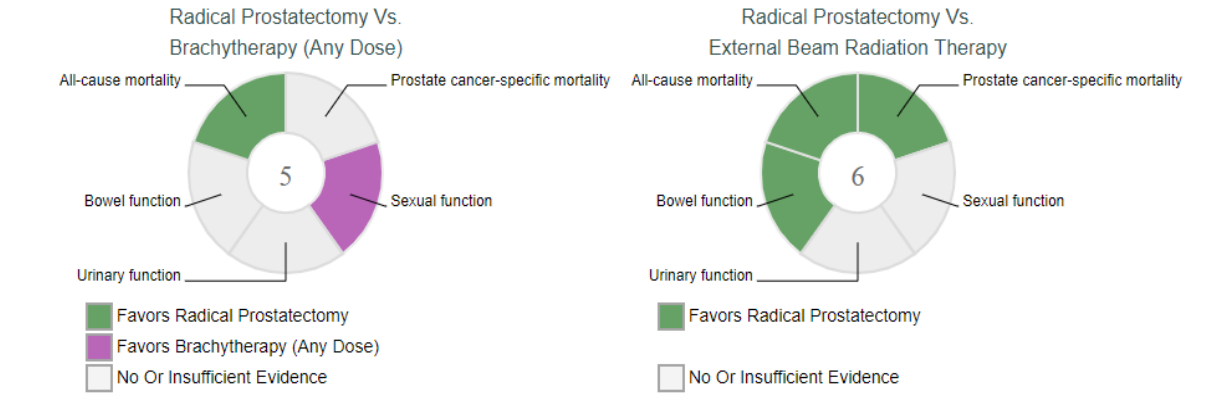
Evidence Visualization 2: Comparison of Treatments by Key Outcome

Evidence Visualization 2 provides detailed assessments of the strength of evidence of each of the 5 outcomes for a selected comparison. Up to 3 comparisons may be reviewed side by side. As depicted in the screenshot in Figure 4, the complete set of outcomes is arrayed in a doughnut. Key visual elements are listed below.

- The **center of the doughnut** lists the number of publications informing the evidence base.
- The **segment color** indicates if the evidence base supports a clear judgment of benefit, harm, no difference, or no evidence.
- The **legend below the doughnut** also clarifies the direction of effect.

- Clicking on any outcome segment in the doughnut expands the page to depict **a panel with the details of the outcome's strength of evidence grade** and provides **a link to citations in PubMed**.

Figure 4. Evidence Visualization 2: Side-by-Side Comparisons of Treatment Comparator Options (with info grid expanded to show details for sexual function as an outcome)



Sexual function

	Radical Prostatectomy Vs. Brachytherapy (Dose Rate Not Specified)	Radical Prostatectomy Vs. External Beam Radiation Therapy
Favors:	Brachytherapy (Any Dose)	Not Applicable
SOE:	Low	Insufficient
Participants: (Publications)	3539 (4)	6353 (6)
Population:	Subpopulation data available	Subpopulation data available
Summary of Findings:	Results consistently favor brachytherapy over radical prostatectomy although the CIs generally span the null.	Confidence intervals of comparison measures do not span the null for 3 of 6 studies and a fourth study is close to the null, suggesting worse sexual function for radical prostatectomy as compared with external beam radiation therapy; the remaining 2 report mixed results for multiple measures (sexual bother and function), all of which span the null.
Consistency:	Consistent (mostly)	Consistent (mostly)
Precision:	Imprecise	Imprecise
Limitations:	Moderate	Moderate
Publications: [Study Design]	Blanchard (2017) [Cohort], Chen (2017) [Cohort], Punnen (2015) [Cohort], Smith (2009) [Cohort]	Barocas (2017) [Cohort], Chen (2017) [Cohort], Punnen (2015) [Cohort], Siegel (2001) [Cohort], Smith (2009) [Cohort], Resnick (2013) [Cohort]
	Link to Studies	Link to Studies

Description of Evidence Base

Table 4 summarizes the characteristics of 40 studies that assessed all-cause or prostate cancer–specific mortality or incidence of sexual, urinary, or bowel harms. Most of these studies were cohort studies (82.5%). A third of the cohort studies drew on patients in the Surveillance, Epidemiology, and End Results (SEER) database who were diagnosed with prostate cancer during the 1990s to early 2000s. Of the 7 noncohort studies, 6 were RCTs (15.0%) and 1 was a non-RCT (2.5%). These studies largely drew on patients in the United States (72.5%), though the remaining studies also drew on patients from countries with very high human development (ie, Australia, Belgium, Canada, Finland, France, Germany, Iceland, Italy, Netherlands, Spain, Sweden, Switzerland, United Kingdom). Follow-up periods ranged from 30 days to 15 years, with most studies (90%) reporting outcomes over follow-up periods longer than 1 year. We determined the quality of a fifth of these studies to be good. Most studies (80.0%) produced some concerns about risk of bias that resulted in a fair rating. We excluded 2 studies that reported on the key outcomes owing to poor quality.[32,33]

Table 5 lists the baseline characteristics for participants in the 40 included studies. Reported age ranges vary, with the most common median age in the 60s. Among studies reporting race (k = 26), most of the study populations were mostly white (ranging from 61.9% to 100%). The average proportion of black study participants was 11.4%. Studies that reported prostate cancer tumor categories (k = 33) reported between 23.6% and 100.0% of participants with T1 disease (mean = 56.5%) and between 13.3% and 76.1% of participants with T2 disease (mean = 41.2%). Most studies reporting baseline Gleason scores (87.5% of 24) reported that more than half of participants had scores of less than 7 compared with less than half with scores of 7 or higher. Among studies reporting prostate cancer risk (k = 24), except one study targeting high-risk disease, most of the study populations had low- or intermediate-risk disease at baseline (ranging from 63.6% to 100.0%). The average proportion of participants with low- or intermediate-risk disease was 79.6%.

Table 4. Study Characteristics

Study first author, publication year	Patient source	Study design	Risk of bias	Country	Follow-up (y)	Total N
Abdollah, 2011[34]	SEER database, 1988-2006	Cohort	Fair	US	Mean 5.1	404 604
Abdollah, 2011[35] **and related publications**[36]	SEER database, 1992-2005	Cohort	Fair/Fair+	US	Planned 10	44 694
Abdollah, 2012[37]	SEER database, 1992-2005	Cohort	Fair	US	Planned 10	68 665
Abdollah, 2012[38]	SEER database, 1992-2005	Cohort	Fair	US	Planned 10	68 797
Albertsen, 2007[39]	Connecticut Tumor Registry	Cohort	Fair	US	Mean 13.3	1352
Albisinni, 2017[40]	NR	Cohort	Fair	Belgium	Median 3	110
Alibhai, 2009[41]	Ontario Cancer Registry	Cohort	Fair–	Canada	Planned 0.1	30 114
Arvold, 2011[42]	NR	Cohort	Fair–	US	Median (by arm) 3.6-6.1	8839
Azzouzi, 2017[43]	NR	RCT	Fair+	Multiple European countries	Planned 2	413
Barocas, 2017[9]	NR	Cohort	Good	US	Planned 3	2550
Bill-Axelson, 2014[44] **and related publications**[45-54]	Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4)	RCT	Good	Multiple European countries	Median 13.4	695
Blanchard, 2018[55]	NR	Cohort	Fair	US	Planned 4	359
Chang, 2017[56]	Prost-QA	Cohort	Fair	US	Planned 2	1021
Chen, 2017[10]	North Carolina Central Cancer Registry	Cohort	Good	US	Planned 2	1141
Cooperberg, 2010[57]	CaPSURE database	Cohort	Fair–	US	Mean (SD) 4.2 (3.3)	6209
Fransson, 2009[58] **and related publications**[59]	Umeå 1	RCT	Fair	Sweden	Median 2.5	108
Giberti, 2009[60]	NR	RCT	Fair+	Italy	Planned 5	200
Hadley, 2010[61]	SEER database, 1995-2007	Cohort	Fair+	US	NR	17 815

Study first author, publication year	Patient source	Study design	Risk of bias	Country	Follow-up (y)	Total N
Hamdy, 2016[62] and related publications[63,64]	Prostate Testing for Cancer and Treatment (ProtecT) Trial	RCT	Good	UK	Median 10	1643
Hansen, 2014[65]	SEER database, 1998-2005	Cohort	Fair+	US	Planned 0.3	59 010
Hoffman, 2003[66] and related publications[67]	SEER database (Prostate Cancer Outcomes Study)	Cohort	Fair	US	Planned 2	2186
Hoffman, 2013[68] and related publications[69]	SEER database (Prostate Cancer Outcomes Study)	Cohort	Good	US	Planned 15	1655
Kibel, 2012[70] and related publications[71]	NR	Cohort	Fair–	US	Planned 10	10 429
Kim, 2011[72]	SEER database, 1992-2005	Cohort	Fair+	US	Planned 10	42 074
Litwin, 1995[73] and related publications[74]	NR	Cohort	Fair	US	Planned 6	214
Litwin, 2004[75]	CaPSURE database	Cohort	Fair+	US	Planned 2	1584
Lu-Yao, 2015[76]	SEER database, 1992-2007	Cohort	Good	US	Median 10.8	57 749
Punnen, 2015[77]	CaPSURE database	Cohort	Good	US	Median 6.2	2273
Rice, 2013[78]	Center for Prostate Disease Research Multicenter National Database	Cohort	Fair	US	Median 6.4	770
Schapira, 2001[79]	NR	Cohort	Fair	US	Planned 1	122
Siegel, 2001[80]	NR	Cohort	Fair	US	Median 4.3	802
Smith, 2000[81]	NR	Cohort	Fair	US	Mean 3.8	1584
Smith, 2009[82]	NR	Cohort	Fair	Australia	Planned 3	1575
Sooriakumaran, 2014[83]	Swedish registries	Cohort	Fair+	Sweden	Range 2.7-8.2	25 197
Stattin, 2010[84]	National Prostate Cancer Register of Sweden Follow-up Study	Cohort	Fair	Sweden	Median 8.2	6849
Thong, 2010[85]	NR	Cohort	Fair	Netherlands	Range 5-10	142
Weissbach, 2016[86] and related publications[87]	Hormonal Therapy, Active Surveillance, Radiation, Operation, Watchful Waiting (HAROW) Study	nRCT	Fair/Fair–	Germany	Median 2.3	2753

Study first author, publication year	Patient source	Study design	Risk of bias	Country	Follow-up (y)	Total N
Westover, 2012[88]	NR	Cohort	Fair–	US	Median 4.6	657
Wilt, 2017[89] **and related publications**[90-93]	Prostate Cancer Intervention Versus Observation Trial (PIVOT)	RCT	Good	US	Median 12.7	731
Wong, 2006[94]	SEER database, 1991-1999	Cohort	Good	US	Planned 12	44 149

Abbreviations: CaPSURE, Cancer of the Prostate Strategic Urologic Research Endeavor; NR, not reported; nRCT, nonrandomized controlled trial; Prost-QA, Prostate Cancer Outcomes and Satisfaction with Treatment Quality Assessment; RCT, randomized controlled trial; SEER, Surveillance, Epidemiology, and End Results; UK, United Kingdom; US, United States.

Table 5. Baseline Characteristics of Study Participants

Study first author, publication year	Age (y)	Race (%)			PSA level (ng/mL)	(%)			Gleason score (%)			Risk (%)		
		White	Black	Other		T1a-c	T2a-c	>T2	<7	7+	Un-known	Low	Inter-mediate	High
Abdollah, 2011[34]	Range 30-95	81.1	11.5	5.0	NR	42.3	57.7	0.0	NR	NR	0.0	67.2[a]		32.8
Abdollah, 2011[35] and related publications[36]	Mean (by arm) 69.8-73.5	85.9	8.5	5.6	NR	42.7	57.3	0.0	NR	NR	0.0	70.9[a]		29.1
Abdollah, 2012[37]	Range 65-80	87.8	7.3	4.9	NR	38.4	61.6	0.0	NR	NR	0.0	63.6[a]		36.4
Abdollah, 2012[38]	Range 65-80	85.4	8.9	5.6	NR	44.3	55.7	0.0	NR	NR	0.0	75.1[a]		24.9
Albertsen, 2007[39]	Median (by arm) 65-71	NR	NR	NR	Median (by arm) 6.6-10.3	NR	NR	NR	59.6	40.4	0.0	32.7	36.0	31.4
Albisinni, 2017[40]	Median (by arm) 63-73	NR	NR	NR	Median (by arm) 6.5-6.9	43.6	56.4	0.0	65.5	34.5	0.0	52.7	40.0	7.3
Alibhai, 2009[41]	Mean (by arm) 62.6-69.1	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Arvold, 2011[42]	Median 67	NR	NR	NR	Median (by arm and risk) 5.3-10.1	75.8	24.2	0.0	80.8	19.2	0.0	65.2	34.8	0.0
Azzouzi, 2017[43]	Mean 63.5	98.8	NR	1.2	Mean 6.1	86.6	13.3	0.0	NR	NR	NR	NR	NR	NR
Barocas, 2017[9]	Mean 63.8	73.5	14.0	24.9	NR	78.5	23.9	NR	52.1	47.9	0.0	45.0	38.6	16.4
Bill-Axelson, 2014[44] and related publications[45-54]	Mean 65	NR	NR	NR	Mean 12.9	23.6	76.1	0.3	60.6	27.9	11.5	35.8	40.4	23.7
Blanchard, 2017[55]	Mean (by arm) 58.7-61.4	81.8	12.8	5.3	Mean (by arm) 5.2-5.4	78.4	21.6	0.0	35.6	64.4	0.0	33.9	64.1	2.0
Chang, 2017[56]	Median 63	NR	8.5[b]	NR	Median 5.5	73.7	26.3	0.0	60.7	39.3	0.0	53.2	37.5	9.3

Study first author, publication year	Age (y)	Race (%)			PSA level (ng/mL)	(%)			Gleason score (%)			Risk (%)		
		White	Black	Other		T1a-c	T2a-c	>T2	<7	7+	Un-known	Low	Inter-mediate	High
Chen, 2017[10]	Median 65	72.3	25.4	2.3	NR	78.6	19.5	1.1	57.23	41.81	1.0	NR	NR	NR
Cooperberg, 2010[57]	Median 65	85.7	10.8	3.4	NR	48.3	49.3	2.4	63.9	36.1	0.0	44.7	34.6	13.6
Fransson, 2009[58] and related publications[59]	Median 77.5	NR	NR	NR	NR	24.6	75.4	0.0	NR	NR	NR	NR	NR	NR
Giberti, 2009[60]	Mean (by arm) 65.2-65.6	100.0	0.0	0.0	Mean (by arm) 7.5-7.8	61.5	38.5	0.0	100.0	0.0	0.0	100.0	0.0	0.0
Hadley, 2010[61]	Range 66-74	77.5	10.6	11.9	NR	63.0	36.4	0.6	NR	NR	NR	NR	NR	NR
Hamdy, 2016[62] and related publications[63,64]	Median 62	97.7	0.6	2.3	Median 5.8	100.0	0.0	0.0	77.1	43.2	0.1	NR	NR	NR
Hansen, 2014[65]	Median 71	81.8	10.5	7.7	NR	39.1	54.8	6.2	NR	NR	NR	NR	NR	NR
Hoffman, 2003[66] and related publications[67]	Mean 66	74.2	13.0	12.8	NR	36.2	40.6	23.2	77.0	22.9	7.7	59.1	NR	NR
Hoffman, 2013[68] and related publications[69]	Median (by arm) 64-69	71.1	14.2	14.7	NR	31.9	44.1	0.0	NR	NR	10.6	45.2	19.2	35.6
Kibel, 2012[70] and related publications[71]	Median (by arm) 60-70	NR	13.0[b]	NR	Median (by arm) 5.2-8.9	73.5	23.9	2.0	70.2	29.8	0.0	56.9	31.3	11.8
Kim, 2011[72]	Range 66-85	79.7	11.7	8.7	NR	56.3	43.7	0.0	NR	NR	3.3	NR	NR	NR
Litwin, 1995[73] and related publications[74]	Mean 73	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Litwin, 2004[75]	Mean 62.8	83.0	5.9	11.2	NR	37.6	49.1	13.4	74.5	20.2	5.3	NR	NR	NR

Study first author, publication year	Age (y)	Race (%)			PSA level (ng/mL)	(%)			Gleason score (%)			Risk (%)		
		White	Black	Other		T1a-c	T2a-c	>T2	<7	7+	Un-known	Low	Inter-mediate	High
Lu-Yao, 2015[76]	Median 74	NR	10.4[b]	NR	Median (by arm) 6.6-6.8	43.4	56.6	0.0	51.1	49.1	0.0	NR	NR	NR
Punnen, 2015[77]	Mean (by arm) 60-72.5	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	58.3	26.1	5.7
Rice, 2013[78]	Mean 74.2	84.4	10.1	5.5	Mean 5.4	60.9	39.1	0.0	100.0	0.0	0.0	100.0	0.0	0.0
Schapira, 2001[79]	Median 68.9	88.5	9.0	2.5	Median (by arm) 7.1-7.9	50.0	50.0	0.0	74.3	13.0	12.3	NR	NR	NR
Siegel, 2001[80]	NR	70.2	23.7	6.1	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Smith, 2000[81]	Mean (by arm) 68-75	NR	3.8[b]	NR	NR	NR	NR	2.0	NR	NR	NR	NR	NR	NR
Smith, 2009[82]	Mean 61.2	NR	NR	NR	Median 7.7	54.4	45.7	0.0	55.0	45.0	0.0	NR	NR	NR
Sooriakumaran, 2014[83]	Range 57-71	NR	NR	NR	Range 4.4-14.0	58.1	38.0	3.8	66.3	33.7	0.0	51.8	48.2	0.0
Stattin, 2010[84]	Mean (by arm) 61.2-64.7	NR	NR	NR	Mean (by arm) 7.6-9.3	58.6	41.4	0.0	85.8	14.2	0.0	39.2	60.8	0.0
Thong, 2010[85]	Mean 76	NR	NR	NR	NR	80.3	19.7	0.0	NR	NR	NR	100.0	0.0	0.0
Weissbach, 2016[86] and related publications[87]	Mean 70.7	NR	NR	NR	NR	66.5	33.6	0.0	62.3	30.2	7.5	45.0	29.6	19.8
Westover, 2012[88]	Median (by arm) 65-70	NR	NR	NR	Median (by arm) 7.9-8.5	58.3	41.7	0.0	0.0	100.0	0.0	0.0	0.0	100.0
Wilt, 2017[89] and related publications[90-93]	Mean 67	61.9	31.8	6.5	Median 7.8	54.7	45.1	0.3	70.5	25.2	NR	40.5	34.1	21.5
Wong, 2006[94]	Mean 72	83.3	8.4	8.4	NR	NR	NR	NR	25.1	74.9	0.0	NR	NR	NR

Abbreviation: NR, not reported; PSA, prostate-specific antigen. [a] Clinical risk reported as low or intermediate vs high. [b] Race reported as black vs other.

Key Findings

Radical Prostatectomy

Fifty-four publications evaluated radical prostatectomy.[9,10,34,35,37,39-42,44,55-57,60-62,65,66,68,70,73,75,77-84,86,88,89,94-114] Twenty-one of these evaluated the outcomes specified in the evidence visualization.[9,10,35,39,41,50,51,53,61-63,66,67,74,77,79,80,82,87,89,91] Six publications compared radical prostatectomy with active surveillance,[9,10,62,63,82,87] a separate set of 6 studies compared radical prostatectomy with watchful waiting,[50,51,53,80,89,91] and 9 compared radical prostatectomy with an unspecified form of conservative management.[35,39,41,61,66,67,74,77,79]

Overall, the strength of the evidence was largely either low or insufficient. Evidence from trials was limited, and evidence from observational studies cannot rule out the potential for confounding. In grading the evidence, we accounted for the potential risk of bias from observational studies. Observational studies that did not adjust for potential confounding were marked as having high study limitations; the strength of evidence was accordingly downgraded. A body of evidence for a particular comparison that entirely comprised observational studies could still yield a strength of evidence of low if the studies attempted to control for confounding and the evidence was consistent, direct, precise, and without evidence of reporting bias.

For the subset of evidence for which we could draw conclusions, the all-cause and prostate cancer–specific mortality evidence favors radical prostatectomy (2 RCTs, low strength of evidence) over watchful waiting, whereas the sexual and urinary physical harms evidence generally favors conservative management strategies (active surveillance, watchful waiting, unspecified conservative management). The evidence for radical prostatectomy vs active surveillance came from 5 studies (1 RCT, 4 observational studies, high strength of evidence) for sexual function and 4 studies (1 RCT, 3 observational studies, high strength of evidence) for urinary function. The evidence for radical prostatectomy vs watchful waiting came from 3 studies (2 RCTs, 1 observational study, low strength of evidence) for sexual function and 2 studies (2 RCTs, moderate strength of evidence) for urinary function. The evidence for radical

prostatectomy vs unspecified conservative management came from 5 observational studies (low strength of evidence) for urinary function. These data trends were consistent across all 3 conservative management comparison groups.

Twenty-seven publications compared radical prostatectomy with radiation therapy for the 5 graded outcomes.[9,10,34,37,39,41,42,55-57,60,65-71,74,77-84] Overall, the strength of evidence for studies comparing these treatments was either low or insufficient based on available evidence. Nonetheless, study results tend to favor radical prostatectomy over unspecified radiation therapy, brachytherapy, and external beam radiation therapy for all-cause and prostate cancer–specific mortality. Specifically, the evidence for radical prostatectomy vs unspecified radiation therapy came from 1 observational study, the evidence for radical prostatectomy vs brachytherapy came from 1 observational study, and the evidence for radical prostatectomy vs external beam radiation therapy came from 6 observational studies. This mortality benefit for radical prostatectomy held for subgroups defined by age, by clinical risk, and by a combination of age and clinical risk; nearly all the evidence came from a single observational study for each subgroup and was graded as low strength of evidence.

Regarding bowel function, the evidence also favors radical prostatectomy over unspecified radiation therapy (2 observational studies, low strength of evidence) and external beam radiation therapy (5 observational studies, low strength of evidence). However, for urinary function, the evidence favors radiation therapy over radical prostatectomy for urinary function, for follow-up periods ranging from 1 to 6 years (3 observational studies, low strength of evidence), and brachytherapy over radical prostatectomy for sexual function, for follow-up periods ranging from 2 to 5 years (4 observational studies, low strength of evidence).

Radiation Therapies

Forty-eight publications evaluated radiation therapies, specifically brachytherapy, external beam radiation therapy, proton therapy, or unspecified radiation therapy.[9,10,34,37-39,41,42,56,57,59,60,65,66,68,70,72,73,75-86,94-100,103,104,106-114] Of these publications, 21 compared radiation therapy with conservative management of various types for the 5 graded

outcomes. [9,10,38,39,41,59,66,67,74,76-82,84,85,94,97,104] We discuss the results of trials (eg, ProtecT) combining external beam radiation therapy with androgen deprivation therapy under the section on combination therapies.

Overall, the evidence is largely insufficient for most outcomes with a few exceptions. Specifically, the evidence may favor radiation therapy (type unspecified) over conservative management for prostate cancer mortality (1 observational study, low strength of evidence) and may favor conservative management over radiation therapy for sexual function (4 observational studies, low strength of evidence) when results are aggregated across subgroups, although the findings for sexual function lack precision and may be the result of chance or confounding.

For unspecified radiation therapy specifically, the prostate cancer–specific mortality benefits when compared with conservative management may hold for subgroups defined by age (aged 70-74 years and aged 75-80 years) and clinical risk; the results lack precision in the 70 to 74 age group. In subgroups defined by baseline Gleason score (in the data file available for download), for those with high Gleason scores, the evidence suggests lower overall mortality and prostate cancer–specific mortality from radiation therapy when compared with conservative management. Among those with low to intermediate Gleason scores, there was no difference in all-cause mortality, but the evidence may suggest benefit for prostate cancer–specific mortality; again, the results lack precision. For subgroups defined by age or Gleason score, the evidence came from a single observational study and yielded low strength of evidence. For subgroups defined by clinical risk, the evidence came from 2 observational studies and also yielded low strength of evidence favoring unspecified radiation therapy over conservative management for all-cause and prostate cancer–specific mortality.

The evidence favors conservative therapy (active surveillance and watchful waiting) over external beam radiation therapy for sexual function in broadly defined populations spanning all risk groups (3 observational studies, low strength of evidence), though not all study results are precise. In low–clinical risk subgroups, the evidence favors active surveillance over external

beam radiation therapy for sexual function and bowel function (1 observational study for each outcome, low strength of evidence) but the evidence is insufficient for other subgroups defined by age, race, or Gleason score. The evidence may favor brachytherapy (any dose) when compared with active surveillance for urinary function (3 observational studies, low strength of evidence); the results lack precision.

Twenty-seven publications reported on comparisons of radiation therapy with radical prostatectomy, described above, in the section on radical prostatectomy. We found no comparative evidence on stereotactic body radiation therapy or proton therapy on the outcomes specified in the evidence visualization. Given the interest in focal treatments to reduce procedure-related harms, further research comparing them against radical prostatectomy and radiation therapy are warranted.

Combination Therapies

Nine publications evaluated combination therapies, specifically radiation therapy plus androgen deprivation therapy and external beam radiation therapy plus brachytherapy.[62,68,72,82,88,98,102,110,111] Of these, 3 evaluated the outcomes specified in the evidence visualization. One publication compared radiation therapy plus androgen deprivation therapy with active surveillance,[82] 2 compared radiation therapy plus androgen deprivation therapy with radical prostatectomy[63,82] and 2 evaluated radical prostatectomy with external beam radiation therapy plus brachytherapy.[102,110]

We did not evaluate studies comparing different types of radiation therapy or comparisons of radiation therapy plus androgen deprivation therapy with androgen deprivation therapy alone or radiation therapy plus androgen deprivation therapy with radiation therapy alone. When comparing combined therapy with active treatment, the results vary depending on the outcome. Specifically, evidence from 2 studies (1 RCT, 1 observational study) suggest improved urinary function for radiation therapy plus androgen deprivation therapy when compared with radical prostatectomy and improved bowel function for radical prostatectomy when compared with radiation therapy plus androgen deprivation therapy. The evidence also

favors radical prostatectomy over external beam radiation therapy plus brachytherapy for bowel function (1 observational study, low strength of evidence). Overall, the evidence is largely insufficient when comparing combination therapies with conservative management. The evidence favors external beam radiation therapy plus brachytherapy over active surveillance for urinary function (1 RCT, low strength of evidence), but it favors active surveillance over radiation therapy plus androgen deprivation therapy for bowel and sexual function (1 observational study each, low strength of evidence) and is insufficient for all other graded outcomes.

Focal Therapies

Three publications evaluated focal therapies, specifically cryosurgery, high-intensity focused ultrasound, and photodynamic therapy.[40,43,81] All 3 evaluated the outcomes specified in the evidence visualization. One publication compared cryosurgery with conservative management,[81] 1 compared high-intensity focused ultrasound with radical prostatectomy,[115] and 1 compared photodynamic therapy with active surveillance.[43]

Overall, the evidence is largely insufficient for most outcomes. For the subset of outcomes for which we can draw conclusions, the evidence does not consistently favor focal therapies, although the strength of evidence is low. Specifically, for sexual function, the evidence may favor high-intensity focused ultrasound (1 observational study, low strength of evidence) over radical prostatectomy and active surveillance over photodynamic therapy (1 RCT, low strength of evidence). The evidence is insufficient for all other graded outcomes. One study reported outcomes for subgroups defined by age for cryosurgery compared with conservative management[81]; the evidence was insufficient to support a judgment of benefit or harm. We found no evidence on irreversible electroporation and MRI-TULSA therapies.

Results by Clinical Risk Factors

The database accompanying the visualization contains detailed assessments of the strength of evidence for specific clinical risk factors when they are available (Table 6). The direction of effect is consistent between the population as a whole and individual subgroups,

although generally weaker in subgroups, often because of higher risk of bias in observational studies or less precise results. Less frequently, a signal of effectiveness could be discerned for the subgroup data (radical prostatectomy vs unspecified radiation therapy, all-cause mortality; unspecified radiation therapy vs unspecified conservative management, all-cause mortality), but not for the population as a whole. For the subgroup, the results were statistically significant and consistent, whereas for the total population, the results were imprecise or inconsistent or lacked adjustment for confounding.

Table 6. Comparing Results for Clinical Risk Subgroups and All Study Participants

Intervention	Comparator	Outcome measure	Subgroups evaluated	Subgroup results	Findings for the total population
Radical prostatectomy	Watchful waiting	Prostate cancer–specific mortality	Age (≥65 years, <65 years)	Low favoring radical prostatectomy	Moderate favoring radical prostatectomy
Radical prostatectomy	Conservative management (unspecified)	Prostate cancer–specific mortality	Age (65-69 years, 70-74 years, 75-80 years), age and clinical risk combinations	Low favoring radical prostatectomy (for aged 65-69, 70-74, 75-80) to insufficient (for aged ≤59 with low or intermediate clinical risk, aged ≤59 with high clinical risk, aged 60-69 with low or intermediate clinical risk, aged 60-69 with high clinical risk, aged 70-79 with low or intermediate clinical risk, aged 70-79 with high clinical risk, aged ≥80 with low or intermediate clinical risk, aged ≥80 with high clinical risk)	Low favoring radical prostatectomy
		Sexual function	Age (≥70 years, <70 years)	Insufficient	Moderate favoring conservative management
		Urinary function	Age (≥70 years, <70 years)	Insufficient	Low favoring conservative management
Radical prostatectomy	Radiation therapy: external beam radiation therapy	All-cause mortality	Clinical risk (low or high), age (55-64 years, 65-70 years, 71-75 years, 65-74 years, >75 years)	Low favoring radical prostatectomy (for low clinical risk, high clinical risk, aged 55-64, aged 55-64) to insufficient (for aged 65-70, 71-75, and >75)	Low favoring radical prostatectomy

Intervention	Comparator	Outcome measure	Subgroups evaluated	Subgroup results	Findings for the total population
		Prostate cancer–specific mortality	Clinical risk (low, intermediate, high), age (55-64 years, 65-74 years)	Low favoring radical prostatectomy (for aged 55-64 and 65-74) to insufficient (for low, intermediate, and high clinical risk)	Low favoring radical prostatectomy
		Sexual function	Race (white, black), Gleason score (2-4, 5-7, 8-10), age (≤59 years, 60-70 years, ≥71 years)	Insufficient	Insufficient
Radical prostatectomy	Radiation therapy: brachytherapy (dose rate not specified)	Prostate cancer–specific mortality	Clinical risk (low, intermediate, high)	Insufficient	Insufficient
		Sexual function	Clinical risk (low)	Insufficient	No studies
		Urinary function	Clinical risk (low)	Insufficient	No studies
Radical prostatectomy	Cryosurgery	Urinary function	Age (≥70 years, <70 years)	Insufficient	Insufficient
		Sexual function	Age (≥70 years, <70 years)	Insufficient	Insufficient
Radiation therapy (type not specified)	Conservative management (unspecified)	All-cause mortality	Clinical risk (low or intermediate), Gleason score (2-6/7, 6/7)	Low favoring radiation therapy (for low or intermediate clinical risk, Gleason score 6/7) to low for no difference (for Gleason 2-6/7)	Insufficient

Intervention	Comparator	Outcome measure	Subgroups evaluated	Subgroup results	Findings for the total population
		Prostate cancer–specific mortality	Clinical risk (low, intermediate, low or intermediate, high), Gleason score (2-6/7, 6/7) age (65-69 years, 70-74 years, 75-80 years)	Low favoring radiation therapy (for low or intermediate clinical risk, high clinical risk, Gleason score 2-6/7, Gleason score 6/7, aged 70-74, aged 75-80) to insufficient (for low clinical risk, intermediate clinical risk, aged 65-69)	Low favoring radiation therapy
		Urinary function	Age (≥70 years, <70 years)	Insufficient	Insufficient
		Sexual function	Age (≥70 years, <70 years)	Insufficient	Low favoring conservative management
Radiation therapy (type not specified)	Cryosurgery	Sexual function	Age (≥70 years, <70 years)	Insufficient	Insufficient
		Urinary function	Age (≥70 years, <70 years)	Insufficient	Insufficient
Radiation therapy: external beam radiation therapy	Watchful waiting	All-cause mortality	Clinical risk (low)	Insufficient	No studies
		Sexual function	Race (white, black), Gleason score (2-4, 5-7), age (≤59 years, 60-70 years, ≥71 years)	Insufficient	Low favoring watchful waiting

Intervention	Comparator	Outcome measure	Subgroups evaluated	Subgroup results	Findings for the total population
Radiation therapy: external beam radiation therapy	Active surveillance	Sexual function	Race (white, Hispanic, black), clinical risk (low)	Low favoring active surveillance (for low clinical risk) to insufficient (for white, Hispanic, black)	Low favoring active surveillance
		Urinary function	Race (white, Hispanic, black), clinical risk (low)	Insufficient	Insufficient
		Bowel function	Race (white, Hispanic, black), clinical risk (low)	Low favoring active surveillance (for low clinical risk) to insufficient (for white, Hispanic, black)	Insufficient
Radiation therapy: brachytherapy (any dose)	Active surveillance	Sexual function	Clinical risk (low)	Insufficient	Insufficient
		Urinary function	Clinical risk (low)	Insufficient	Low favoring brachytherapy
		Bowel function	Clinical risk (low)	Insufficient	Insufficient
Cryosurgery	Conservative management (unspecified)	Sexual function	Age (≥70 years, <70 years)	Insufficient	Insufficient
		Urinary function	Age (≥70 years, <70 years)	Insufficient	Insufficient

Limitations

This work has important limitations. First, as noted above, the exclusion of within-class comparisons limits the ability of these evidence visualizations to address the comparative effectiveness of all relevant treatment options for clinically localized prostate cancer. We exclude, for example, studies comparing radiation therapy plus androgen deprivation therapy with androgen deprivation therapy alone or with radiation therapy alone. These restrictions on eligible comparisons mean that our conclusions regarding the effectiveness of combination therapy are very specific to included comparators and cannot be interpreted more generally regarding the effectiveness of combination therapy as a whole. Second, the Heat Map and Evidence Visualization 2 are restricted to studies that reported on the 5 main outcomes in the evidence visualization and do not comment on their time horizon. Studies report on other outcomes for some comparisons not depicted in the Heat Map and Evidence Visualization 2. Third, the evidence visualization does not depict the entire array of factors influencing the choice of a treatment. Resolving the uncertainty around the choice of a treatment plan will require information on other outcomes and other considerations not depicted in the evidence visualization. Fourth, we generally did not combine across treatment categories (except brachytherapy). As a result, although the heterogeneity within each comparison group was reduced, we could not offer comment on broad comparisons, such as radical prostatectomy vs any conservative management strategy. Fifth, our visualization choices to manage the complexity of the information may have resulted in some oversimplification. For example, the grades we assigned are not consistent across all age and clinical risk subgroups and the visualization currently does not include a filter to depict these variations. Similarly, although strength-of-evidence grades incorporate the risk of confounding in observational studies, the visualization currently does not include a filter for study design so that users can easily identify evidence bases primarily comprising RCTs. Finally, our reliance on existing systematic reviews as the base for the update of the evidence may have introduced some inconsistencies in the eligible evidence. Although we re-reviewed inclusion, exclusion, and risk-of-bias decisions from

prior reports, it is possible that underlying inconsistencies resulted in the exclusion of potentially relevant evidence.

Conclusions

Systematic reviews are rich sources of evidence that support medical decision making, especially for preference-sensitive conditions like early-stage prostate cancer. Our evidence review update confirms that benefits of active treatments also carry significant harms compared with conservative approaches. This pilot project advances the use of visual tools to unpack the complex information of a systematic review, translating tables of data into accessible visualizations of the amount and strength of evidence for different treatments comparisons. These evidence visualizations provide a broad overview of the comparative effectiveness of early-stage prostate cancer therapies for a set of key outcomes, while providing users access to the rich database of evidence underpinning the visuals. This pilot project offers a transparent approach of how to map systematic review report data to specific visual elements, translating complex data into interactive visual evidence products for future researchers seeking to improve dissemination of systematic reviews.

References

1. Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin.* 2016;66(4):271-289.

2. National Cancer Institute. Cancer stat facts: prostate cancer. 2018; https://seer.cancer.gov/statfacts/html/prost.html. Accessed February 19, 2018.

3. Negoita S, Feuer EJ, Mariotto A, et al. Annual Report to the Nation on the Status of Cancer, part II: Recent changes in prostate cancer trends and disease characteristics. *Cancer.* 2018;124(13):2801-2814.

4. Force USPST, Grossman DC, Curry SJ, et al. Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA.* 2018;319(18):1901-1913.

5. Bell KJ, Del Mar C, Wright G, Dickinson J, Glasziou P. Prevalence of incidental prostate cancer: a systematic review of autopsy studies. *Int J Cancer.* 2015;137(7):1749-1757.

6. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin.* 2017;67(1):7-30.

7. Bruner DW, Moore D, Parlanti A, Dorgan J, Engstrom P. Relative risk of prostate cancer for men with affected relatives: systematic review and meta-analysis. *Int J Cancer.* 2003;107(5):797-803.

8. Fenton JJ, Weyrich MS, Durbin S, Liu Y, Bang H, Melnikow J. Prostate-specific antigen-based screening for prostate cancer: evidence report and systematic review for the US Preventive Services Task Force. *JAMA.* 2018;319(18):1914-1931.

9. Barocas DA, Alvarez J, Resnick MJ, et al. Association between radiation therapy, surgery, or observation for localized prostate cancer and patient-reported outcomes after 3 years. *JAMA.* 2017;317(11):1126-1140.

10. Chen RC, Basak R, Meyer AM, et al. Association between choice of radical prostatectomy, external beam radiotherapy, brachytherapy, or active surveillance and patient-reported quality of life among men with localized prostate cancer. *JAMA.* 2017;317(11):1141-1150.

11. Litwin MS, Tan HJ. The Diagnosis and Treatment of Prostate Cancer: A Review. *JAMA.* 2017;317(24):2532-2542.

12. Valerio M, Cerantola Y, Eggener SE, et al. New and Established Technology in Focal Ablation of the Prostate: A Systematic Review. *Eur Urol.* 2017;71(1):17-34.

13. Patient-Centered Outcomes Research Institute. A prospective comparative study of outcomes with proton and photon radiation in prostate cancer. 2017;

https://www.pcori.org/research-results/2017/prospective-comparative-study-outcomes-proton-and-photon-radiation-prostate. Accessed February 6, 2017.

14. Fenton JJ, Weyrich MS, Durbin S, Liu Y, Bang H, Melnikow J. Prostate-specific antigen–based screening for prostate cancer: A systematic evidence review for the U.S. Preventive Services Task Force. Sacramento, CA: Center for Healthcare Policy and Research; 2017.

15. American Urological Association. Clinically localized prostate cancer: AUA/ASTRO/SUO Guideline. 2017; http://www.auanet.org/guidelines/clinically-localized-prostate-cancer-new-(aua/astro/suo-guideline-2017).

16. Fenton JJ, Weyrich MS, Durbin S, Liu Y, Bang H, Melnikow J. *Prostate-specific antigen-based screening for prostate cancer: a systematic evidence review for the U.S. Preventive Services Task Force.* Rockville, MD: Evidence Synthesis, No. 154; May 2018. Report No.: 17-05229-EF-1.

17. Sun F, Oyesanmi O, Fontanarosa J, Reston J, Guzzon T, Schoelles K. *Therapies for clinically localized prostate cancer. Update of a 2008 systematic review.* Rockville, MD: Comparative Effectiveness Reviews, No. 146; December 2014. Report No.: 15-EHC004-EF.

18. . *Prostate Cancer: Diagnosis and Treatment*. Cardiff (UK)2014.

19. Carter HB, Albertsen PC, Barry MJ, et al. Early detection of prostate cancer: AUA Guideline. *J Urol.* 2013;190(2):419-426.

20. Sanda MG, Cadeddu JA, Kirkby E, et al. Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part I: Risk Stratification, Shared Decision Making, and Care Options. *J Urol.* 2018;199(3):683-690.

21. National Comprehensive Cancer Network. Prostate Cancer (Version 3.2016). 2016; https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf.

22. Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. *Eur Urol.* 2016;69(3):428-435.

23. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol.* 2011;64(12):1283-1293.

24. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2011;64(4):383-394.

25. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol.* 2011;64(4):401-406.

26. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol.* 2011;64(4):407-415.

27. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol.* 2011;64(12):1277-1282.

28. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol.* 2011;64(12):1294-1302.

29. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol.* 2011;64(12):1303-1310.

30. Guyatt GH, Oxman AD, Sultan S, et al. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol.* 2011;64(12):1311-1316.

31. Berkman ND, Lohr KN, Ansari MT, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. *J Clin Epidemiol.* 2015;68(11):1312-1324.

32. Carlsson S, Drevin L, Loeb S, et al. Population-based study of long-term functional outcomes after prostate cancer treatment. *BJU Int.* 2016;117(6b):E36-45.

33. Gielchinsky I, Chang J, Cusick T, et al. Prostate cancer in 432 men aged <50 years in the prostate-specific antigen era: a new outlook. *BJU Int.* 2018;122 Suppl 5:35-41.

34. Abdollah F, Sun M, Thuret R, et al. A competing-risks analysis of survival after alternative treatment modalities for prostate cancer patients: 1988-2006. *Eur Urol.* 2011;59(1):88-95.

35. Abdollah F, Sun M, Schmitges J, et al. Cancer-specific and other-cause mortality after radical prostatectomy versus observation in patients with prostate cancer: competing-risks analysis of a large North American population-based cohort. *Eur Urol.* 2011;60(5):920-930.

36. Abdollah F, Sun M, Schmitges J, et al. Survival benefit of radical prostatectomy in patients with localized prostate cancer: estimations of the number needed to treat according to tumor and patient characteristics. *J Urol.* 2012;188(1):73-83.

37. Abdollah F, Schmitges J, Sun M, et al. Comparison of mortality outcomes after radical prostatectomy versus radiotherapy in patients with localized prostate cancer: a population-based analysis. *International journal of urology : official journal of the Japanese Urological Association.* 2012;19(9):836-844.

38. Abdollah F, Sun M, Schmitges J, et al. Competing-risks mortality after radiotherapy vs. observation for localized prostate cancer: a population-based study. *Int J Radiat Oncol Biol Phys.* 2012;84(1):95-103.

39. Albertsen PC, Hanley JA, Penson DF, Barrows G, Fine J. 13-year outcomes following treatment for clinically localized prostate cancer in a population based cohort. *J Urol.* 2007;177(3):932-936.

40. Albisinni S, Aoun F, Bellucci S, et al. Comparing high-intensity focal ultrasound hemiablation to robotic radical prostatectomy in the management of unilateral prostate cancer: a matched-pair analysis. *J Endourol.* 2017;31(1):14-19.

41. Alibhai SM, Leach M, Warde P. Major 30-day complications after radical radiotherapy: a population-based analysis and comparison with surgery. *Cancer.* 2009;115(2):293-302.

42. Arvold ND, Chen MH, Moul JW, et al. Risk of death from prostate cancer after radical prostatectomy or brachytherapy in men with low or intermediate risk disease. *J Urol.* 2011;186(1):91-96.

43. Azzouzi AR, Vincendeau S, Barret E, et al. Padeliporfin vascular-targeted photodynamic therapy versus active surveillance in men with low-risk prostate cancer (CLIN1001 PCM301): an open-label, phase 3, randomised controlled trial. *Lancet Oncol.* 2017;18(2):181-191.

44. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in early prostate cancer. *N Engl J Med.* 2014;370(10):932-942.

45. Bill-Axelson A, Holmberg L, Filen F, et al. Radical prostatectomy versus watchful waiting in localized prostate cancer: the Scandinavian prostate cancer group-4 randomized trial. *J Natl Cancer Inst.* 2008;100(16):1144-1154.

46. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med.* 2011;364(18):1708-1717.

47. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med.* 2005;352(19):1977-1984.

48. Holmberg L, Bill-Axelson A, Garmo H, et al. Prognostic markers under watchful waiting and radical prostatectomy. *Hematol Oncol Clin North Am.* 2006;20(4):845-855.

49. Holmberg L, Bill-Axelson A, Helgesen F, et al. A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer. *N Engl J Med.* 2002;347(11):781-789.

50. Johansson E, Bill-Axelson A, Holmberg L, Onelov E, Johansson JE, Steineck G. Time, symptom burden, androgen deprivation, and self-assessed quality of life after radical prostatectomy or watchful waiting: the Randomized Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4) clinical trial. *Eur Urol.* 2009;55(2):422-430.

51. Johansson E, Steineck G, Holmberg L, et al. Long-term quality-of-life outcomes after radical prostatectomy or watchful waiting: the Scandinavian Prostate Cancer Group-4 randomised trial. *Lancet Oncol.* 2011;12(9):891-899.

52. Bill-Axelson A, Garmo H, Holmberg L, et al. Long-term distress after radical prostatectomy versus watchful waiting in prostate cancer: a longitudinal study from the Scandinavian Prostate Cancer Group-4 randomized clinical trial. *Eur Urol.* 2013;64(6):920-928.

53. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical Prostatectomy or Watchful Waiting in Prostate Cancer - 29-Year Follow-up. *N Engl J Med.* 2018;379(24):2319-2329.

54. Holmberg L, Bill-Axelson A, Steineck G, et al. Results from the Scandinavian Prostate Cancer Group Trial Number 4: a randomized controlled trial of radical prostatectomy versus watchful waiting. *J Natl Cancer Inst Monogr.* 2012;2012(45):230-233.

55. Blanchard P, Davis JW, Frank SJ, et al. Quality of life after brachytherapy or bilateral nerve-sparing robot-assisted radical prostatectomy for prostate cancer: a prospective cohort. *BJU Int.* 2018;121(4):540-548.

56. Chang P, Regan MM, Ferrer M, et al. Relief of Urinary Symptom Burden after Primary Prostate Cancer Treatment. *J Urol.* 2017;197(2):376-384.

57. Cooperberg MR, Vickers AJ, Broering JM, Carroll PR. Comparative risk-adjusted mortality outcomes after primary surgery, radiotherapy, or androgen-deprivation therapy for localized prostate cancer. *Cancer.* 2010;116(22):5226-5234.

58. Fransson P, Damber JE, Widmark A. Health-related quality of life 10 years after external beam radiotherapy or watchful waiting in patients with localized prostate cancer. *Scand J Urol Nephrol.* 2009;43(2):119-126.

59. Fransson P, Damber JE, Tomic R, Modig H, Nyberg G, Widmark A. Quality of life and symptoms in a randomized trial of radiotherapy versus deferred treatment of localized prostate carcinoma. *Cancer.* 2001;92(12):3111-3119.

60. Giberti C, Chiono L, Gallo F, Schenone M, Gastaldi E. Radical retropubic prostatectomy versus brachytherapy for low-risk prostatic cancer: a prospective study. *World J Urol.* 2009;27(5):607-612.

61. Hadley J, Yabroff KR, Barrett MJ, Penson DF, Saigal CS, Potosky AL. Comparative effectiveness of prostate cancer treatments: evaluating statistical adjustments for confounding in observational data. *J Natl Cancer Inst.* 2010;102(23):1780-1793.

62. Hamdy FC, Donovan JL, Lane JA, et al. 10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. *N Engl J Med.* 2016;375(15):1415-1424.

63. Donovan JL, Hamdy FC, Lane JA, et al. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. *N Engl J Med.* 2016;375(15):1425-1437.

64. Lane JA, Donovan JL, Davis M, et al. Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. *Lancet Oncol.* 2014;15(10):1109-1118.

65. Hansen J, Gandaglia G, Bianchi M, et al. Re-assessment of 30-, 60- and 90-day mortality rates in non-metastatic prostate cancer patients treated either with radical prostatectomy or radiation therapy. *Canadian Urological Association journal = Journal de l'Association des urologues du Canada.* 2014;8(1-2):E75-80.

66. Hoffman RM, Hunt WC, Gilliland FD, Stephenson RA, Potosky AL. Patient satisfaction with treatment decisions for clinically localized prostate carcinoma. Results from the Prostate Cancer Outcomes Study. *Cancer.* 2003;97(7):1653-1662.

67. Hoffman RM, Lo M, Clark JA, et al. Treatment decision regret among long-term survivors of localized prostate cancer: results from the Prostate Cancer Outcomes Study. *J Clin Oncol.* 2017;35(20):2306-2314.

68. Hoffman RM, Koyama T, Fan K-H, et al. Mortality after radical prostatectomy or external beam radiotherapy for localized prostate cancer. *J Natl Cancer Inst.* 2013;105(10):711-718.

69. Resnick MJ, Koyama T, Fan K-H, et al. Long-term functional outcomes after treatment for localized prostate cancer. *The New England journal of medicine.* 2013;368(5):436-445.

70. Kibel AS, Ciezki JP, Klein EA, et al. Survival among men with clinically localized prostate cancer treated with radical prostatectomy or radiation therapy in the prostate specific antigen era. *The Journal of urology.* 2012;187(4):1259-1265.

71. Nepple KG, Stephenson AJ, Kallogjeri D, et al. Mortality after prostate cancer treatment with radical prostatectomy, external-beam radiation therapy, or brachytherapy in men without comorbidity. *Eur Urol.* 2013;64(3):372-378.

72. Kim S, Shen S, Moore DF, et al. Late gastrointestinal toxicities following radiation therapy for prostate cancer. *Eur Urol.* 2011;60(5):908-916.

73. Litwin MS. Health-related quality of life after treatment for localized prostate cancer. *Cancer.* 1995;75(S7):2000-2003.

74. Litwin MS, Hays RD, Fink A, et al. Quality-of-life outcomes in men treated for localized prostate cancer. *JAMA.* 1995;273(2):129-135.

75. Litwin MS, Sadetsky N, Pasta DJ, Lubeck DP. Bowel function and bother after treatment for early stage prostate cancer: a longitudinal quality of life analysis from CaPSURE. *The Journal of urology.* 2004;172(2):515-519.

76. Lu-Yao GL, Kim S, Moore DF, et al. Primary radiotherapy vs conservative management for localized prostate cancer--a population-based study. *Prostate Cancer Prostatic Dis.* 2015;18(4):317-324.

77. Punnen S, Cowan JE, Chan JM, Carroll PR, Cooperberg MR. Long-term health-related quality of life after primary treatment for localized prostate cancer: results from the CaPSURE registry. *Eur Urol.* 2015;68(4):600-608.

78. Rice KR, Colombo ML, Wingate J, et al. Low risk prostate cancer in men $\geq$ 70 years old: to treat or not to treat. *Urol Oncol.* 2013;31(6):755-760.

79. Schapira MM, Lawrence WF, Katz DA, McAuliffe TL, Nattinger AB. Effect of treatment on quality of life among men with clinically localized prostate cancer. *Med Care.* 2001;39(3):243-253.

80. Siegel T, Moul JW, Spevak M, Alvord WG, Costabile RA. The development of erectile dysfunction in men treated for prostate cancer. *J Urol.* 2001;165(2):430-435.

81. Smith DS, Carvalhal GF, Schneider K, Krygiel J, Yan Y, Catalona WJ. Quality-of-life outcomes for men with prostate carcinoma detected by screening. *Cancer.* 2000;88(6):1454-1463.

82. Smith DP, King MT, Egger S, et al. Quality of life three years after diagnosis of localised prostate cancer: population based cohort study. *BMJ.* 2009;339:b4817.

83. Sooriakumaran P, Nyberg T, Akre O, et al. Comparative effectiveness of radical prostatectomy and radiotherapy in prostate cancer: observational study of mortality outcomes. *BMJ.* 2014;348:g1502.

84. Stattin P, Holmberg E, Johansson JE, Holmberg L, Adolfsson J, Hugosson J. Outcomes in localized prostate cancer: National Prostate Cancer Register of Sweden follow-up study. *J Natl Cancer Inst.* 2010;102(13):950-958.

85. Thong MS, Mols F, Kil PJ, Korfage IJ, van de Poll-Franse LV. Prostate cancer survivors who would be eligible for active surveillance but were either treated with radiotherapy

or managed expectantly: comparisons on long-term quality of life and symptom burden. *BJU Int.* 2010;105(5):652-658.

86. Weissbach L, Stuerzebecher S, Mumperow E, Klotz T, Schnell D. HAROW: the first comprehensive prospective observational study comparing treatment options in localized prostate cancer. *World J Urol.* 2016;34(5):641-647.

87. Hilger C, Schostak M, Neubauer S, et al. The importance of sexuality, changes in erectile functioning and its association with self-esteem in men with localized prostate cancer: data from an observational study. *BMC Urol.* 2019;19(1):9.

88. Westover K, Chen MH, Moul J, et al. Radical prostatectomy vs radiation therapy and androgen-suppression therapy in high-risk prostate cancer. *BJU Int.* 2012;110(8):1116-1121.

89. Wilt TJ, Jones KM, Barry MJ, et al. Follow-up of prostatectomy versus observation for early prostate cancer. *N Engl J Med.* 2017;377(2):132-142.

90. Wilt TJ. The Prostate Cancer Intervention Versus Observation Trial: VA/NCI/AHRQ Cooperative Studies Program #407 (PIVOT): design and baseline results of a randomized controlled trial comparing radical prostatectomy with watchful waiting for men with clinically localized prostate cancer. *J Natl Cancer Inst Monogr.* 2012;2012(45):184-190.

91. Wilt TJ, Brawer MK, Jones KM, et al. Radical Prostatectomy versus Observation for Localized Prostate Cancer. *N Engl J Med.* 2012;367(3):203-213.

92. Wilt TJ, Brawer MK, Barry MJ, et al. The Prostate cancer Intervention Versus Observation Trial:VA/NCI/AHRQ Cooperative Studies Program #407 (PIVOT): design and baseline results of a randomized controlled trial comparing radical prostatectomy to watchful waiting for men with clinically localized prostate cancer. *Contemp Clin Trials.* 2009;30(1):81-87.

93. Wilt TJ. Management of low risk and low PSA prostate cancer: long term results from the prostate cancer intervention versus observation trial. *Recent Results Cancer Res.* 2014;202:149-169.

94. Wong YN, Mitra N, Hudes G, et al. Survival associated with treatment vs observation of localized prostate cancer in elderly men. *JAMA.* 2006;296(22):2683-2693.

95. Jang JW, Drumm MR, Efstathiou JA, et al. Long-term quality of life after definitive treatment for prostate cancer: patient-reported outcomes in the second posttreatment decade. *Cancer Med.* 2017;6(7):1827-1836.

96. Williams SB, Huo J, Chamie K, et al. Discerning the survival advantage among patients with prostate cancer who undergo radical prostatectomy or radiotherapy: The limitations of cancer registry data. *Cancer.* 2017;123(9):1617-1624.

97. Tyson MD, Alvarez J, Koyama T, et al. Racial Variation in Patient-Reported Outcomes Following Treatment for Localized Prostate Cancer: Results from the CEASAR Study. *Eur Urol.* 2017;72(2):307-314.

98. Williams SB, Duan Z, Chamie K, et al. Risk of hospitalisation after primary treatment for prostate cancer. *BJU Int.* 2017;120(1):48-55.

99. Wallis CJ, Mahar AL, Satkunasivam R, et al. Cardiovascular and Skeletal-related Events Following Localized Prostate Cancer Treatment: Role of Surgery, Radiotherapy, and Androgen Deprivation. *Urology.* 2016;97:145-152.

100. Barbosa PV, Thomas IC, Srinivas S, et al. Overall survival in patients with localized prostate cancer in the US Veterans Health Administration: is PIVOT generalizable? *Eur Urol.* 2016;70(2):227-230.

101. Basu A, Gore JL. Are Elderly Patients With Clinically Localized Prostate Cancer Overtreated? Exploring Heterogeneity in Survival Effects. *Med Care.* 2015;53(1):79-86.

102. Hjalm-Eriksson M, Lennernas B, Ullen A, et al. Long-term health-related quality of life after curative treatment for prostate cancer: a regional cross-sectional comparison of two standard treatment modalities. *Int J Oncol.* 2015;46(1):381-388.

103. Ravi P, Karakiewicz PI, Roghmann F, et al. Mental health outcomes in elderly men with prostate cancer. *Urol Oncol.* 2014;32(8):1333-1340.

104. Acar C, Schoffelmeer CC, Tillier C, de Blok W, van Muilekom E, van der Poel HG. Quality of life in patients with low-risk prostate cancer. A comparative retrospective study: brachytherapy versus robot-assisted laparoscopic prostatectomy versus active surveillance. *J Endourol.* 2014;28(1):117-124.

105. Jeldres C, Cullen J, Hurwitz LM, et al. Prospective quality-of-life outcomes for low-risk prostate cancer: Active surveillance versus radical prostatectomy. *Cancer.* 2015;121(14):2465-2473.

106. Schymura MJ, Kahn AR, German RR, et al. Factors associated with initial treatment and survival for clinically localized prostate cancer: results from the CDC-NPCR Patterns of Care Study (PoC1). *BMC Cancer.* 2010;10:152.

107. Nam RK, Cheung P, Herschorn S, et al. Incidence of complications other than urinary incontinence or erectile dysfunction after radical prostatectomy or radiotherapy for

prostate cancer: a population-based cohort study. *The Lancet Oncology.* 2014;15(2):223-231.

108. Lughezzani G, Sun M, Perrotte P, et al. Comparative study of inguinal hernia repair rates after radical prostatectomy or external beam radiotherapy. *Int J Radiat Oncol Biol Phys.* 2010;78(5):1307-1313.

109. Ladjevardi S, Sandblom G, Berglund A, Varenhorst E. Tumour grade, treatment, and relative survival in a population-based cohort of men with potentially curable prostate cancer. *Eur Urol.* 2010;57(4):631-638.

110. Huang GJ, Sadetsky N, Penson DF. Health related quality of life for men treated for localized prostate cancer with long-term followup. *J Urol.* 2010;183(6):2206-2212.

111. Zhou EH, Ellis RJ, Cherullo E, et al. Radiotherapy and survival in prostate cancer patients: a population-based study. *Int J Radiat Oncol Biol Phys.* 2009;73(1):15-23.

112. Lubeck DP, Litwin MS, Henning JM, Stoddard ML, Flanders SC, Carroll PR. Changes in health-related quality of life in the first year after treatment for prostate cancer: results from CaPSURE. *Urology.* 1999;53(1):180-186.

113. Bacon CG, Giovannucci E, Testa M, Kawachi I. The impact of cancer treatment on quality of life outcomes for patients with localized prostate cancer. *J Urol.* 2001;166(5):1804-1810.

114. Ferrer M, Guedea F, Suárez JF, et al. Quality of life impact of treatments for localized prostate cancer: cohort study with a 5 year follow-up. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology.* 2013;108(2):306-313.

115. Umbehr MH, Bachmann LM, Poyet C, et al. The German version of the Expanded Prostate Cancer Index Composite (EPIC): translation, validation and minimal important difference estimation. *Health Qual Life Outcomes.* 2018;16(1):36.

Appendix 1: Search Strategy

Appendix 1 Table 1. PubMed Search String and Yield (11/27/2019)

Search	Query	Items found
#1	("Prostate"[Mesh]) OR "Prostatic Neoplasms"[Mesh]	135 981
#2	(cancer*[tiab] or carcinoma*[tiab] or neoplasm*[tiab])	2 096 820
#3	prostat*[tiab]	198 788
#4	(#2 and #3)	138 662
#5	("prostate cancer"[tiab] OR "prostatic neoplasm*"[tiab])	102 968
#6	(#1 or #4 or #5)	176 589
#7	("watchful waiting"" [tiab] OR "active surveillance"" tiab] OR LRP[tiab] OR RLRP[tiab] OR prostatectom*[tiab] OR radiotherap*[tiab] OR EBRT[tiab] OR IMRT[tiab] OR proton[tiab] OR (intensity[tiab] AND modulated[tiab] AND therap*[tiab]) OR brachytherap*[tiab] OR curietherap*[tiab] OR cryosurger*[tiab] OR cryotherap*[tiab] OR cryoablat*[tiab] OR Cyberknife[tiab] OR freezing[tiab] OR HIFU[tiab] OR (high[tiab] AND intensity[tiab] AND focused[tiab] AND ultrasound*[tiab]))	356 272
#8	("Watchful Waiting"[Mesh] OR "Brachytherapy"[Mesh] OR "Cryosurgery"[Mesh] OR "Cryotherapy"[Mesh] OR "Freezing"[Mesh] OR "High-Intensity Focused Ultrasound Ablation"[Mesh] OR "Prostatectomy"[Mesh] OR "Radiotherapy"[Mesh])	260 047
#9	"therapy" [Subheading] OR "Treatment Outcome"[Mesh] OR "Therapeutics"[Mesh]	8 335 082
#10	(#7 or #8 or #9)	8 493 210
#11	(("Early Diagnosis"[Mesh] OR early stage[tw]) or ("Early Diagnosis"[tiab] OR "early stage"[tiab]))	181 994
#12	"Early Detection of Cancer"[Mesh] OR "Early Medical Intervention"[Mesh]	21 950
#13	("early stage*"[tiab] or "early detection"[tiab] or "localized"[tiab] or "localized"[tiab] or "early intervention"[tiab])	408 919

Search	Query	Items found
#14	t1[tiab] or t2[tiab] or “stage 1”[tiab] or “stage one”[tiab] or “nonmetastatic”[tiab] or “non-metastatic”[tiab]	144 703
#15	#11 or #12 or #13 or #14	639 164
#16	(#6 and #10 and #15)	15 093
#17	(#6 and #10 and #15) Filters: Publication date from 2014/01/01	3978
#18	(#6 and #10 and #15) Filters: Publication date from 2014/01/01; Humans	3128

Appendix 1 Table 2. Cochrane Search String and Yield (11/27/2019)

Search	Query	Items found
#1	Prostate cancer prostat* AND (neoplasm* OR cancer* OR carcinoma*)	12 007
#2	Treatment options "watchful waiting" OR "active surveillance" OR LRP OR RLRP OR prostatectom* OR radiotherap* OR EBRT OR IMRT OR proton OR (intensity AND modulated AND therap*) OR brachytherap* OR curietherap* OR cryosurger* OR cryotherap* OR cryoablat* OR Cyberknife OR freezing OR HIFU OR (high AND intensity AND focused AND ultrasound*)	35 651
#3	Combine sets #1 AND #2	44 147
#4	Limit 3 to: Publication date from 2014 to 11/2018	57

Appendix 1 Table 3. AHRQ Evidence Reports and Technology Assessments Search String and Yield (1/7/2019)

Search	Query	Items found
#1	"Prostate Cancer"; Filters: Publication date from 2014/01/01	11

Abbreviation: AHRQ, Agency for Healthcare Research and Quality.

Appendix 1 Table 4. Drug Effectiveness Review Project Drug Class Reviews Search String and Yield (1/7/2019)

Search	Query	Items found
#1	Ctrl+F "prostate"	0
#2	Ctrl+F "cancer"	0

Appendix 1 Table 5. National Guidelines Search String and Yield (1/7/2019)

Search	Query	Items found
#1	"Prostate Cancer"; Filters: Publication date from 2014/01/01	12

Appendix 1 Table 6. ClinicalTrials.gov Search String and Yield (1/10/2019)

Search	Query	Items found
#1	"Localized Prostate Cancer" AND "Treatment"; Filters: Primary completion from 2014/01/01; Recruiting, not yet recruiting, active not recruiting, or enrolling by invitation	29

Appendix 1 Table 7. HSRPproj Search String and Yield (1/10/2019)

Search	Query	Items found
#1	"Localized Prostate Cancer" AND "Treatment"; Filters: Ongoing	3

Appendix 1 Table 8. PCORI Portfolio Search String and Yield (1/10/2019)

Search	Query	Items found
#1	Filters: Cancer; Prostate cancer; Ongoing or PCORI peer review	9

Appendix 2: Risk-of-Bias Assessment

Appendix 2 Table 1. Guidance for Risk of Bias Assessment for Randomized Controlled Trials

Field	Instructions
Valid random assignment?	Yes, No, Uncertain, NR If researchers used random number table and no suggestion of irregularities, random assignment was likely valid. Use of terms like "computer-generated," "random number table," or "block randomization" suggests valid random assignment. Code "No" if the method of assignment may be related to the outcome in some way or there is other cause for concern. Code "Uncertain" for methods that may not be truly random (eg, those based on characteristics of the person, setting, or date, but not likely to be correlated with the outcome).
Allocation concealment	Yes, No, Uncertain, NR Code "Yes" if the allocation process appears to be impervious to any influence by the individual making the allocation. Inadequate approaches (code as "No") include alternation, the use of case record numbers, dates of birth or day of the week, and any procedure that is transparent before allocation (such as an open list of random numbers).
Groups similar at baseline?	Yes; No, adjusted; No; Uncertain; NR Code "Yes" if all intervention and control groups are similar in age, sex, ethnicity, education, socioeconomic status, or any other baseline characteristics affecting prognosis. Code "No, adjusted" if there were baseline differences but they are adjusted for in the analysis.
Eligibility criteria specified?	Yes, No, Uncertain, NR
Measurements: equal, reliable, valid	Likely, No, Uncertain, NR Code "Likely" if instruments were previously developed and references are provided (unless there are known issues with the instrument) or if reliability/validity information is provided AND if procedures are identical for all treatment groups, unless there are other issues of concern.
Blinding of (1) participants,	For each, code Yes, No, Uncertain, NR, N/A May not need to include blinding of participants and providers for

Field	Instructions
(2) providers, (3) outcomes assessors	some topics, such as counseling topics, for which blinding of participants and providers is essentially impossible. Code "Yes" if blinding of outcomes assessors is explicitly stated.
Differential follow-up?	Report the percentage with follow-up at XX months for all treatment arms included in the review, if available. If not available, choose the next closest follow-up that is at least YY months post-baseline. If 2 follow-up assessments are equidistant from XX months, choose the one that occurs after the XXth month. *Topics must specify values of XX and YY.*
Intention-to-treat analysis?	ITT, Completers only, Per protocol only, No Code as "ITT" if at least XX% (98%) of randomized participants are included in the analysis. If data substitution methods were used, such as multiple imputation, mean substitution, or last observation carried forward, describe briefly. Code as "Completers only" if only those providing some or all follow-up data are included in the analysis (and this is less than 98% of randomized participants). Code "Per protocol" if only those with some specified level of exposure to the intervention are included in the analysis. Code "No" if participants were not analyzed in the group to which they were allocated.
Intervention fidelity?	Yes, No, Not assessed, Uncertain, NR Code "Yes" if researchers provide a clear description of the intervention and some effort at quality control of intervention delivery was described, such as the use of scripts and manuals, supervision, or observation.
Adequate adherence to treatment?	Yes, No, Not assessed, Uncertain, NR Code "Yes" if some measure of exposure to intervention is reported and is judged to be adequate (eg, attended 80% of sessions)
Funding source	Government, Industry, Other, NR
Other important threats to validity (not already captured)	This field will be carried forward to the summary table. Any threats to validity that are not captured by the columns that will be in the summary table should be listed here. Reviewers should together agree on what should be included in this column.
Quality rating	Good, Fair+, Fair, Fair–, Poor
Comments	Narrative summary of quality issues. Note whether any issues could be considered fatal flaws. Clarify anything coded as "Uncertain."

Abbreviations: ITT, intention to treat; N/A, not applicable; NR, not reported.

Appendix 2 Table 2. Guidance for Risk-of-Bias Assessment for Observational Studies

Field	Instructions
Selection of cohort	Is the cohort systematically selected to ensure avoidance of bias? Appropriate, Inappropriate, Uncertain, NR
Selection of the nonexposed cohort	Was the nonexposed cohort systematically selected? Appropriate, Inappropriate, Uncertain, NR
Ascertainment of exposure	Was the ascertainment of exposure reported? Yes, No, Uncertain, NR
Groups similar at baseline?	Yes; No, adjusted; No; Uncertain; NR Code "Yes" if all intervention and control groups are similar in age, sex, ethnicity, education, socioeconomic status, or any other baseline characteristics affecting prognosis. Code "No, adjusted" if there were baseline differences but they are adjusted for in the analysis.
Eligibility criteria specified?	Yes, No, Uncertain, NR
Outcome of interest is not present at baseline	Was the outcome of interest not present at baseline? Yes, No, Uncertain, NR
Measurements: equal, reliable, valid	Likely, No, Uncertain, NR Code "Likely" if instruments were previously developed and references are provided (unless there are known issues with the instrument), or if reliability/validity information is provided, or if assessment is very straightforward and no real cause for concern AND if procedures are identical for all treatment groups, unless there are other issues of concern. Code "No" if procedures or assessment personnel are different for different treatment groups, or other causes for concern.
Blinding of outcomes assessors	Yes, No, Uncertain, NR Code "Yes" if any one of the following: No blinding, but the review authors judge that the outcome and the outcome measurement are not likely to be influenced by lack of blinding; Blinding of participants and key study personnel was ensured, and unlikely that the blinding could have been broken; Either participants or some key study personnel were not blinded, but outcome assessment was blinded and the nonblinding of others is unlikely to introduce bias.

Field	Instructions
	Code “No” if any one of the following: No blinding or incomplete blinding, and the outcome or outcome measurement is likely to be influenced by lack of blinding; Blinding of key study participants and personnel was attempted, but likely that the blinding could have been broken; Either participants or some key study personnel were not blinded, and the nonblinding of others is likely to introduce bias
Acceptable length of follow-up	Was follow-up long enough for the outcome to occur? Likely, No, Uncertain, NR
Acceptable follow-up	Excellent, Adequate, Problematic, NR Code “Excellent” if 90% follow-up or higher overall, not differential. Code “Adequate” if 60% to 89% follow-up, not differential and any one of the following: Reasons for missing outcome data are unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias); Missing outcome data are balanced in numbers across intervention groups, with similar reasons for missing data across groups; For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk is not enough to have a clinically relevant impact on the intervention effect estimate; For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes is not enough to have a clinically relevant impact on observed effect size; Missing data have been imputed using appropriate methods. Code “Problematic” if follow-up < 60%, or does not meet “adequate” criteria, or differential follow-up (rules of thumb: >20% differential follow-up is poor, 10%-20% differential follow-up is questionable).
Adjustment for potential confounders	Adjustment for all important factors, Adjustment for some of the important factors, No adjustment
Funding source	Government, Industry, Other, NR

Field	Instructions
Other important threats to validity (not already captured)	This field will be carried forward to the summary table. Any threats to validity that are not captured by the columns that will be in the summary table should be listed here. Reviewers should together agree on what should be included in this column.
Quality rating	Good, Fair+, Fair, Fair–, Poor
Comments	Narrative summary of quality issues. Note whether any issues could be considered fatal flaws. Clarify anything coded as "Uncertain."

Abbreviations: NR, not reported.