HHS Public Access

Author manuscript

Eur Urol. Author manuscript; available in PMC 2018 December 01.

Published in final edited form as:

Eur Urol. 2017 December ; 72(6): 899–907. doi:10.1016/j.eururo.2017.07.018.

Active Surveillance Versus Watchful Waiting for Localized Prostate Cancer: A Model to Inform Decisions

Stacy Loeb[a,b,*], **Qinlian Zhou**[b], **Uwe Siebert**[d,e,f,g], **Ursula Rochau**[d], **Beate Jahn**[d], **Nikolai Mühlberger**[d], **H. Ballentine Carter**[h], **Herbert Lepor**[a], and **R. Scott Braithwaite**[b]

[a]Department of Urology, New York University, New York, NY, USA

[b]Department of Population Health, New York University, New York, NY, USA

[c]The Manhattan Veterans Affairs Medical Center, New York, NY, USA

[d]Department of Public Health, Health Services Research and Health Technology Assessment, UMIT—University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria

[e]ONCOTYROL—Center for Personalized Cancer Medicine, Innsbruck, Austria

[f]Center for Health Decision Science and Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[g]Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[h]The Brady Institute of Urology, Johns Hopkins Medical Institutions, Baltimore, MD, USA

Abstract

[*]Corresponding author. 550 1st Ave VZ30 (6th Floor, 612), New York, NY 10016, USA. Tel. +1 (646) 501-2559; Fax: +1 (212) 263-4549. stacyloeb@gmail.com (S. Loeb).

Author contributions: Stacy Loeb had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Loeb, Carter, Lepor, Braithwaite.
Acquisition of data: Loeb, Zhou, Mühlberger, Carter, Lepor, Braithwaite.
Analysis and interpretation of data: Loeb, Zhou, Siebert, Rochau, Jahn, Mühlberger, Braithwaite.
Drafting of the manuscript: Loeb, Zhou.
Critical revision of the manuscript for important intellectual content: Loeb, Zhou, Siebert, Rochau, Jahn, Mühlberger, Carter, Lepor, Braithwaite.
Statistical analysis: Zhou, Siebert, Rochau, Jahn, Mühlberger, Braithwaite.
Obtaining funding: Loeb.
Administrative, technical, or material support: Loeb, Braithwaite.
Supervision: Siebert, Carter, Lepor, Braithwaite.
Other: None.

Financial disclosures: Stacy Loeb certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Stacy Loeb received honoraria for lectures from MDxHealth and Boehringer Ingelheim, travel reimbursement from Minomic and Boehringer Ingelheim, and consulting for Lilly (unrelated to the current manuscript).

Background—An increasing proportion of prostate cancer is being managed conservatively. However, there are no randomized trials or consensus regarding the optimal follow-up strategy.

Objective—To compare life expectancy and quality of life between watchful waiting (WW) versus different strategies of active surveillance (AS).

Design, setting, and participants—A Markov model was created for US men starting at age 50, diagnosed with localized prostate cancer who chose conservative management by WW or AS using different testing protocols (prostate-specific antigen every 3–6 mo, biopsy every 1–5 yr, or magnetic resonance imaging based). Transition probabilities and utilities were obtained from the literature.

Outcome measurements and statistical analysis—Primary outcomes were life years and quality-adjusted life years (QALYs). Secondary outcomes include radical treatment, metastasis, and prostate cancer death.

Results and limitations—All AS strategies yielded more life years compared with WW. Lifetime risks of prostate cancer death and metastasis were, respectively, 5.42% and 6.40% with AS versus 8.72% and 10.30% with WW. AS yielded more QALYs than WW except in cohorts age >65 yr at diagnosis, or when treatment-related complications were long term. The preferred follow-up strategy was also sensitive to whether people value short-term over long-term benefits (time preference). Depending on the AS protocol, 30–41% underwent radical treatment within 10 yr. Extending the surveillance biopsy interval from 1 to 5 yr reduced life years slightly, with a 0.26 difference in QALYs.

Conclusions—AS extends life more than WW, particularly for men with higher-risk features, but this is partly offset by the decrement in quality of life since many men eventually receive treatment.

Patient summary—More intensive active surveillance protocols extend life more than watchful waiting, but this is partly offset by decrements in quality of life from subsequent treatment.

Keywords

Prostate cancer; Active surveillance; Watchful waiting; Conservative management; Markov model

1. Introduction

Prostate cancer (PCa) screening reduces advanced disease and PCa-specific death [1,2], but also leads to “overdiagnosis” and overtreatment of indolent tumors [3,4]. Conservative management is increasingly utilized for favorable-risk PCa to delay or avoid aggressive treatment and potential side effects [5]. Prior comparative-effectiveness models have confirmed that this is a valid strategy for certain patients [6–8], with improved quality of life (QOL) and reduced initial resource utilization [9].

Despite agreement on the importance of conservative management to preserve screening benefits and reduce overtreatment [10], there is no consensus what to do next [11,12]. Conservative management encompasses two very different strategies: “watchful waiting” (WW) without curative intent and “active surveillance” (AS) with serial testing for “disease

progression” to offer selective delayed treatment with curative intent. No randomized trials have compared benefits and harms between WW and contemporary AS. Furthermore, for patients choosing AS, there is no consensus on the type, frequency, or sequence of follow-up tests to monitor for disease progression [11]. Thus, the objective of this clinical decision analysis is to compare life expectancy and quality-adjusted life expectancy between WW and different AS protocols for US men ≥50 yr.

2. Patients and methods

We developed a state-transition Markov model to compare different strategies of conservative management for a cohort of US men diagnosed with clinically localized PCa who chose conservative management. Markov models represent a hypothetical cohort moving among predefined health states that are mutually exclusive and collectively exhaustive [13]. Our model starts when the patient is diagnosed with PCa and begins conservative management. We used this model to evaluate two different outcomes: life years (LYs) and quality-adjusted life years (QALYs), which put quality and quantity of life into the same metric by multiplying the predicted duration of each health state by the utility (QOL weight) for living in that state. The model was analyzed and reported according to ISPOR/SMDM international recommendations [13].

The base case analyses compare WW (follow without further testing until the development of advanced PCa or death from other causes) with AS with prostate-specific antigen (PSA) every 6 mo and yearly biopsy (based on the Johns Hopkins AS protocol [14]). We also examined an AS strategy with more frequent PSAs (quarterly) with biopsies at years 1, 3, 7, and 10, and then every 5 yr, similar to Prostate Cancer Research International Active Surveillance (PRIAS) [15], and an exploratory strategy including PSA every 6 mo and magnetic resonance imaging (MRI) yearly where biopsy is performed only if MRI is abnormal. Finally, we evaluated an exploratory strategy with PSA every 6 mo and biopsy every 5 yr. For all strategies, biopsies were discontinued at age 75 yr in the main analysis, as in the Johns Hopkins program [14].

We used a state-transition cohort model to obtain estimates for specific populations of interest determined a priori, based on clinical features. For the main analysis, the cohort started at age 50 yr, and the model was rerun for cohorts starting at age 40, 65, 70, and 75 yr. Figure 1 shows a schematic of the model. At the start, men have been diagnosed with PCa and they have chosen conservative management. Some were classified accurately with Gleason 6 (grade group 1), while others were misclassified and have undetected higher-grade disease. During each model cycle, individuals can remain on conservative management, undergo treatment for reclassification (then into a post-treatment state), develop metastases, or die. We used a cycle length of 1 mo and a lifelong time horizon due to the long natural history of PCa. Depending on the approach to conservative management, some cycles may include rebiopsy. Overall mortality data were obtained from US life tables, with a priori adjustment by a multiplier of 0.45 to account for the highly selected healthier population affected by localized PCa [14]. Our model considered the following potential harms: biopsy complications, short- and long-term complications of PCa treatment (aggregate measure including sexual, urinary, and bowel dysfunction), and development of

metastasis. Since our objective was to examine efficacy, we assumed 100% compliance with protocol-recommended biopsies and that all men found to have disease reclassification (increases in tumor grade) underwent treatment.

Table 1 shows the model inputs (see Supplementary material for details). Transition probabilities between states were estimated from the literature. Previously published "utilities" (ie, QOL weights reflecting quantitative health preferences) were used to quantify QOL implications for each disease state [16].

One- and two-way deterministic sensitivity analyses were performed to assess the implications of uncertainty for key variables. Tornado diagrams were used to summarize results of one-way sensitivity analysis. Since previous studies showed an impact of time preference on PCa treatment selection, we also performed sensitivity analysis using discounting (ie, assigning lower weights to future events) [17]. We also estimated the risk of radical treatment, metastasis, and PCa death. Model validation was performed based on ISPOR–SMDM recommendations and comprised the following: (1) expert consensus on face validity of model inputs, structure, and results; (2) verification through extensive sensitivity and extreme value analysis; (3) cross validation to previous models; and (4) blinded external validation to partially dependent and independent published studies with >5 yr follow-up [18]. All analyses were performed using TreeAge Pro version 2014 (TreeAge Software, Inc., Williamstown, MA, USA).

3. Results

3.1. Main base case analysis

Table 2 shows the base case results of the decision analysis. In a cohort of men starting at age 50 with low-risk PCa undergoing conservative management, AS using the Johns Hopkins strategy yielded more LYs compared with WW (35.21 vs 34.55 LYs, or a difference of 0.66 life-years; Table 2). Lifetime risks of PCa death and metastasis were, respectively, 5.42% and 6.40% with AS versus 8.72% and 10.30% with WW. Men on AS had a 50% lifetime risk of undergoing radical treatment.

Using the outcome of quality-adjusted life expectancy, AS yielded more QALYs (33.89) than WW (33.36 QALYs, or a difference of 0.53 life-years).

For a cohort starting at age 40 yr (Table 2), AS yielded more LYs and QALYs compared with WW. By contrast, among men aged ≥65, WW had more QALYs than AS (Table 2). Supplementary Table 1 shows LYs and QALYs for men with very low–risk PCa.

3.2. Alternative AS protocols

In men aged ≥50 yr, using PRIAS, MRI-based, and 5-yr biopsy strategies yielded 35.12, 35.20, and 34.99 LYs, respectively. Lifetime risks of PCa death and metastasis were 6.01% and 7.10% with PRIAS, 5.40% and 6.39% with the MRI-based, and 6.93% and 8.19% with the 5-yr biopsy strategies, respectively. Lifetime risks of receiving radical treatment were 46% with PRIAS, 50% with the MRI-based, and 43% with 5-yr biopsy strategies. AS using the PRIAS (33.79 QALYs), MRI-based (33.89 QALYs), and 5-yr biopsy (33.63 QALYs)

strategies yielded higher QALYs than WW (33.36 QALYs). Supplementary Table 2 shows the 10-yr and lifetime risks of receiving radical treatment, metastasis, and PCa death for cohorts starting at different ages.

3.3. Sensitivity analyses

In one-way sensitivity analysis for the end point of LYs (Fig. 2A), the risk of metastasis for untreated grade reclassification, age at initiation, proportion with initial misclassification, and ratio of reduction in metastasis with treatment versus WW had the greatest overall impact on remaining life expectancy. However, only age >77.6 yr at initiation led to a switch in preferred strategy from AS to WW based on the outcome of LYs. To test whether this age sensitivity was an artifact of discontinuing biopsies at 75 yr in the base case scenario, we generated a separate model and performed an analysis starting at age 70 and 75 yr with biopsies extending until 85 yr, and AS yielded slightly more LYs (+0.14 and +0.07, respectively).

Given the variability in reported rates of treatment-related complications and difficulties in estimating joint-state utilities for side effects, sensitivity analyses were also performed for the outcome QALYs (Fig. 2B). The discount rate, risk of metastasis for untreated grade reclassification, duration of treatment complications, and age at initiation all had a substantial impact on expected QALY. However, the only parameters leading to a shift in the preferred management strategy were discount rate (0.0018), risk of metastasis (2.4% at 10y), duration of treatment-related complications (>27 y) and age.

In two-way sensitivity analyses (Fig. 3A), AS was preferred with a shorter duration and lower decrement in utility from treatment-related complications, whereas WW was associated with more QALYs with long-term larger utility decrement from treatment complications. AS was also associated with more QALYs across the range of utilities, except at a very low probability of metastasis and high decrement in utility from treatment-related complications (Fig. 3B).

Two way sensitivity analyses were also performed for discount rates to characterize the substantial sensitivity to time preference (Fig. 3C). Setting the probability of treatment-related complications at 0, WW was the preferred choice if the discount rate is >0.0023 (below the lower plausible bound of 0.03). Even at the highest probability of metastasis used in our sensitivity analysis of 0.002 (or 21% at 10 yr), discounting >0.005 made WW preferred to AS.

4. Discussion

AS extends life more than WW, particularly for men with higher-risk disease with a greater risk of metastasis. However, intensive follow-up protocols with frequent rebiopsy and use of radical treatment for men with grade reclassification may reduce QOL. Extending the interval between biopsies up to 5 yr led fewer men to receive radical treatment, with a small reduction in incremental LYs and QALYs. Time preferences and duration of QOL decrements from treatment side effects also had a significant impact on the results. These findings show the importance of shared decision making; these trade-offs should be

discussed with patients to provide decisions regarding the intent and intensity of conservative management options based on individualized patient preferences [19].

These results are particularly timely given recent evidence that the use of conservative management for PCa is rapidly increasing. In the USA, there was a significant spike in the use of conservative management up to >40% in 2010–2013 [20], with similar trends internationally [21]. Nationwide Swedish data showed that 91% of very low–risk and 74% of low-risk patients chose AS in 2014 [22].

Despite increasing utilization, there are limited data determining what to do next for men choosing conservative management and real-world practice patterns vary widely [23,24]. A 2011 National Institutes of Health (NIH) consensus conference concluded that “follow-up under AS is variable and not currently evidence-based. The types of monitoring and their optimal frequency need to be defined. It is important to consider whether follow-up should vary based on tumor and patient characteristics" [11]. First, there is no level 1 evidence that AS is superior to WW. Moreover, for patients choosing AS, there is no consensus on the type, frequency, or sequence of follow-up tests to monitor for disease progression [11]. That notwithstanding, the choice of follow-up testing may have significant implications for patients and healthcare system. PSA and digital rectal examination are less invasive and costly, but may not reliably identify disease progression [25]. In a randomized trial, men with screen-detected PCa monitored primarily based on PSA kinetics without regularly scheduled biopsies had a higher risk of metastasis at 10-yr than those who received prostatectomy or radiation therapy [26]. Contemporary AS programs also incorporate serial prostate biopsies every 1–5 yr. Prostate biopsies provide information on grade and tumor volume [27], but are invasive with increasing infectious complications [28]. Finally, numerous AS programs have recently begun using MRI, reporting a high positive predictive value for disease progression [29–31]. MRI is more expensive and time consuming than blood or urinary markers, but less invasive than biopsy. No data from prospective, randomized trials are published comparing alternative conservative management strategies.

Decision-analytic modeling studies are useful in such situations with multiple management alternatives with substantial tradeoffs and no randomized evidence supporting one approach over another [32]. The results of our decision analysis provide novel data demonstrating that the testing regimen during AS has only a small impact on estimates of LYs or QALYs. By contrast, tumor features, treatment-related morbidity, and patient preferences may have a large impact on the preferred approach to conservative management, suggesting that patient-shared decision making with an individualized assessment of tumor characteristics and patient preferences is important even once a patient has chosen to defer treatment [19]. Although there are challenges associated with performing preference assessment in clinical practice, this is an area of significant active investigation [33].

Randomized trials comparing surgery versus observation suggested that certain subgroups have greater benefit from aggressive treatment (eg, age <65 yr, PSA >10) [8,34]. Our model suggests that some of these same factors also affect the preferred approach to conservative management, with trade-offs between more intensive testing to detect reclassification in time for curative treatment with potential side effects, versus less intensive testing without

curative intent. These results are consistent with what has been observed comparing various AS approaches in the literature [35]. Factors that increase the risk of severe, lasting treatment-related complications also favor a less intensive approach to conservative management (WW), whereas factors increasing the risk of metastasis with untreated cancer favor a more intensive approach (AS). Overall, there was limited benefit to performing additional biopsies after age 75 yr (<10 yr life expectancy) and already more harm than benefit in the cohort aged ≥65 yr, suggesting that a transition to WW around this time is reasonable.

We also observed that the preferred choice of monitoring during conservative management was exquisitely sensitive to time preference. A recent study found that time discounting was negatively associated with choice of prostatectomy over AS [17]. Our results expand upon this for men who have already chosen conservative management, showing that a high discount rate (focused on well-being in the present) favors WW, while a low discount rate (places more emphasis on future well-being) favors AS. This raises the question as to whether time preferences should be assessed as part of clinical decision-making; however, this presents logistical challenges and would require further research on how to perform such an assessment in clinical practice.

As in all decision analyses, this study has several limitations, including uncertainty for several model parameters. We performed extensive sensitivity analysis to make this uncertainty transparent, revealing that few parameters had a substantive impact on model results. Notable exceptions are the extent and duration of QOL impact from treatment-related complications, which vary widely in the literature [36]. That notwithstanding, our model suggested that a switch in the preferred decision would only occur with severe, lasting treatment-related complications. Another drawback to our study is limited published data for many AS testing strategies. However, the model suggests that the precise protocol is not among the key determinants of LYs or QALYs, confirming the robustness of our results. Similarly, the amount of initial misclassification may be lower using new genomic markers and MRI-targeted biopsy. However, the results were robust, and inferences for decision making changed neither in sensitivity analyses with a hypothetical scenario of 0% initial misclassification, nor in sensitivity analysis improving MRI performance characteristics, suggesting that these are also not key determinants of LYs or QALYs. Another limitation is that we used a Markov cohort simulation, precluding the ability to track test results over time. Follow-up studies using microsimulation [19] are warranted given that reclassification is a conditional probability [37]. Finally, in order to compare the efficacy of different protocols under ideal conditions, we assumed 100% compliance with protocol-indicated biopsies and treatment recommendations when reclassification occurs. While we begin to incorporate these data into patient counseling, future studies are warranted, including an effectiveness analysis with real-world adherence rates and incorporating other end points such as cost effectiveness, which are also critical for healthcare decision making [38].

5. Conclusions

AS extends life more than WW, but this is partly offset by the decrement in QOL since a substantial proportion ultimately undergo radical treatment. Patient preferences had a

significant influence on model results, and further research is warranted on how to optimally incorporate preference assessment into clinical practice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

***Funding/Support and role of the sponsor*:** This study was supported by the Edward Blank and Sharon Cosloy-Blank Family Foundation, the Gertrude and Louis Feil Family, the New York State Department of Health (DOH01-C30697GG-3450000), the Laura and Isaac Perlmutter Cancer Center at NYU Langone Medical Center (P30CA016087), and the National Institutes of Health (Award Number K07CA178258) to Stacy Loeb. The content is solely the responsibility of the authors and does not represent the official views of the NIH.

References

1. Schroder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. N Engl J Med. 2012; 366:981–90. [PubMed: 22417251]
2. Etzioni R, Tsodikov A, Mariotto A, et al. Quantifying the role of PSA screening in the US prostate cancer mortality decline. Cancer Causes Control. 2008; 19:175–81. [PubMed: 18027095]
3. Barry MJ. Screening for prostate cancer--the controversy that refuses to die. N Engl J Med. 2009; 360:1351–4. [PubMed: 19297564]
4. Etzioni R, Penson DF, Legler JM, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst. 2002; 94:981–90. [PubMed: 12096083]
5. Gulati R, Wever EM, Tsodikov A, et al. What if I don't treat my PSA-detected prostate cancer? Answers from three natural history models. Cancer Epidemiol Biomarkers Prev. 2011; 20:740–50. [PubMed: 21546365]
6. Hayes JH, Ollendorf DA, Pearson SD, et al. Active surveillance compared with initial treatment for men with low-risk prostate cancer: a decision analysis. JAMA. 2010; 304:2373–80. [PubMed: 21119084]
7. Liu D, Lehmann HP, Frick KD, Carter HB. Active surveillance versus surgery for low risk prostate cancer: a clinical decision analysis. J Urol. 2012; 187:1241–6. [PubMed: 22335873]
8. Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med. 2012; 367:203–13. [PubMed: 22808955]
9. Kim S 2nd, Dall'Era MA, Evans CP. Economic analysis of active surveillance for localized prostate cancer. Curr Opin Urol. 2012; 22:47–53. [PubMed: 22080871]
10. Roth JA, Gulati R, Gore JL, Cooperberg MR, Etzioni R. Economic analysis of prostate-specific antigen screening and selective treatment strategies. JAMA Oncol. 2016; 2:890–8. [PubMed: 27010943]
11. Ganz PA, Barry JM, Burke W, et al. National Institutes of Health State-of-the-Science Conference: role of active surveillance in the management of men with localized prostate cancer. Ann Intern Med. 2012; 156:591–5. [PubMed: 22351514]
12. Barry MJ. The prostate cancer treatment bazaar: comment on "Physician visits prior to treatment for clinically localized prostate cancer". Arch Intern Med. 2010; 170:450–2. [PubMed: 20212181]
13. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR–SMDM Modeling Good Research Practices Task Force-3. Med Decis Making. 2012; 32:690–700. [PubMed: 22990084]
14. Tosoian JJ, Mamawala M, Epstein JI, et al. Intermediate and longer-term outcomes from a prospective active-surveillance program for favorable-risk prostate cancer. J Clin Oncol. 2015; 33:3379–85. [PubMed: 26324359]
15. Bul M, Zhu X, Valdagni R, et al. Active surveillance for low-risk prostate cancer worldwide: the PRIAS study. Eur Urol. 2013; 63:597–603. [PubMed: 23159452]

16. Heijnsdijk EA, Wever EM, Auvinen A, et al. Quality-of-life effects of prostate-specific antigen screening. N Engl J Med. 2012; 367:595–605. [PubMed: 22894572]
17. Lopez-Perez, B., Barnes, A., Frosch, DL., Hanoch, Y. Predicting prostate cancer treatment choices: the role of numeracy, time discounting, and risk attitudes. J Health Psychol. In press. http://dx.doi.org/10.1177/1359105315615931
18. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR–SMDM Modeling Good Research Practices Task Force-7. Med Decis Making. 2012; 32:733–43. [PubMed: 22990088]
19. Rogowski W, Payne K, Schnell-Inderst P, et al. Concepts of 'personalization' in personalized medicine: implications for economic evaluation. Pharmacoeconomics. 2015; 33:49–59. [PubMed: 25249200]
20. Cooperberg MR, Carroll PR. Trends in management for patients with localized prostate cancer, 1990–2013. JAMA. 2015; 314:80–2. [PubMed: 26151271]
21. Loeb S, Berglund A, Stattin P. Population based study of use and determinants of active surveillance and watchful waiting for low and intermediate risk prostate cancer. J Urol. 2013; 190:1742–9. [PubMed: 23727309]
22. Loeb, S., Folkvaljon, Y., Curnyn, C., Robinson, D., Bratt, O., Stattin, P. Uptake of active surveillance for very-low-risk prostate cancer in Sweden. JAMA Oncol. In press. http://dx.doi.org/10.1001/jamaoncol.2016.3600
23. Loeb S, Carter HB, Schwartz M, Fagerlin A, Braithwaite RS, Lepor H. Heterogeneity in active surveillance protocols worldwide. Rev Urol. 2014; 16:202–3. [PubMed: 25548550]
24. Loeb S, Curnyn C, Fagerlin A, et al. Qualitative study on decision-making by prostate cancer physicians during active surveillance. BJU Int. 2017; 120:32–9. [PubMed: 27611479]
25. Ross AE, Loeb S, Landis P, et al. Prostate-specific antigen kinetics during follow-up are an unreliable trigger for intervention in a prostate cancer surveillance program. J Clin Oncol. 2010; 28:2810–6. [PubMed: 20439642]
26. 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; 375:1415–24. [PubMed: 27626136]
27. Tosoian JJ, Trock BJ, Landis P, et al. Active surveillance program for prostate cancer: an update of the Johns Hopkins experience. J Clin Oncol. 2011; 29:2185–90. [PubMed: 21464416]
28. Loeb S, Carter HB, Berndt SI, Ricker W, Schaeffer EM. Complications after prostate biopsy: data from SEER-Medicare. J Urol. 2011; 186:1830–4. [PubMed: 21944136]
29. Margel D, Yap SA, Lawrentschuk N, et al. Impact of multiparametric endorectal coil prostate magnetic resonance imaging on disease reclassification among active surveillance candidates: a prospective cohort study. J Urol. 2012; 187:1247–52. [PubMed: 22335871]
30. Morgan VA, Riches SF, Thomas K, et al. Diffusion-weighted magnetic resonance imaging for monitoring prostate cancer progression in patients managed by active surveillance. Br J Radiol. 2011; 84:31–7. [PubMed: 21172965]
31. Vargas HA, Akin O, Afaq A, et al. Magnetic resonance imaging for predicting prostate biopsy findings in patients considered for active surveillance of clinically low risk prostate cancer. J Urol. 2012; 188:1732–8. [PubMed: 23017866]
32. Loeb S, Carlsson S, Braithwaite RS. Prostate cancer: modeling the outcomes of prostate cancer screening. Nat Rev Urol. 2012; 9:183–5. [PubMed: 22410678]
33. Shirk JD, Saigal CS. From QOL to QALYs: comparing nononcologic outcomes in prostate cancer survivors across treatments. Urol Oncol. 2017; 35:69–75. [PubMed: 27575917]
34. Bill-Axelson A, Holmberg L, Garmo H, Rider JR, Taari K, Busch C, Nordling S, Häggman M, Andersson SO, Spångberg A, Andrén O, Palmgren J, Steineck G, Adami HO, Johansson JE. Radical prostatectomy or watchful waiting in early prostate cancer. N Engl J Med. 2014 Mar 6; 370(10):932–42. DOI: 10.1056/NEJMoa1311593 [PubMed: 24597866]
35. Tosoian JJ, Carter HB, Lepor A, Loeb S. Active surveillance for prostate cancer: current evidence and contemporary state of practice. Nat Rev Urol. 2016; 13:205–15. [PubMed: 26954332]

36. Dale W, Bilir SP, Hemmerich J, Basu A, Elstein A, Meltzer D. The prevalence, correlates, and impact of logically inconsistent preferences in utility assessments for joint health states in prostate cancer. Med Care. 2011; 49:59–66. [PubMed: 21150801]
37. Alam R, Carter HB, Landis P, Epstein JI, Mamawala M. Conditional probability of reclassification in an active surveillance program for prostate cancer. J Urol. 2015; 193:1950–5. [PubMed: 25572035]
38. Siebert U. When should decision-analytic modeling be used in the economic evaluation of health care? Eur J Health Econ. 2003; 4:143–50.
39. Shapiro RH, Johnstone PA. Risk of Gleason grade inaccuracies in prostate cancer patients eligible for active surveillance. Urology. 2012; 80:661–6. [PubMed: 22925240]
40. Inoue LY, Trock BJ, Partin AW, Carter HB, Etzioni R. Modeling grade progression in an active surveillance study. Stat Med. 2014; 33:930–9. [PubMed: 24123208]
41. Jain S, Loblaw A, Vesprini D, et al. Gleason upgrading with time in a large prostate cancer active surveillance cohort. J Urol. 2015; 194:79–84. [PubMed: 25660208]
42. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in early prostate cancer. N Engl J Med. 2014; 370:932–42. [PubMed: 24597866]
43. Klotz L, Vesprini D, Sethukavalan P, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. J Clin Oncol. 2015; 33:272–7. [PubMed: 25512465]
44. Musunuru HB, Yamamoto T, Klotz L, et al. Active surveillance for intermediate risk prostate cancer: survival outcomes in the Sunnybrook experience. J Urol. 2016; 196:1651–8. [PubMed: 27569437]
45. Popiolek M, Rider JR, Andren O, et al. Natural history of early, localized prostate cancer: a final report from three decades of follow-up. Eur Urol. 2013; 63:428–35. [PubMed: 23084329]
46. Eisenberger MA, Blumenstein BA, Crawford ED, et al. Bilateral orchiectomy with or without flutamide for metastatic prostate cancer. N Engl J Med. 1998; 339:1036–42. [PubMed: 9761805]
47. Sweeney CJ, Chen YH, Carducci M, et al. Chemohormonal therapy in metastatic hormone-sensitive prostate cancer. N Engl J Med. 2015; 373:737–46. [PubMed: 26244877]
48. James ND, Sydes MR, Clarke NW, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. Lancet. 2016; 387:1163–77. [PubMed: 26719232]
49. Gravis G, Fizazi K, Joly F, et al. Androgen-deprivation therapy alone or with docetaxel in non-castrate metastatic prostate cancer (GETUG-AFU 15): a randomised, open-label, phase 3 trial. Lancet Oncol. 2013; 14:149–58. [PubMed: 23306100]
50. Guo R, Cai L, Fan Y, Jin J, Zhou L, Zhang K. Magnetic resonance imaging on disease reclassification among active surveillance candidates with low-risk prostate cancer: a diagnostic meta-analysis. Prostate Cancer Prostatic Dis. 2015; 18:221–8. [PubMed: 25986915]
51. Siddiqui MM, Rais-Bahrami S, Turkbey B, et al. Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. JAMA. 2015; 313:390–7. [PubMed: 25626035]
52. de Rooij M, Hamoen EH, Futterer JJ, Barentsz JO, Rovers MM. Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis. AJR Am J Roentgenol. 2014; 202:343–51. [PubMed: 24450675]
53. Loeb S, van den Heuvel S, Zhu X, Bangma CH, Schroder FH, Roobol MJ. Infectious complications and hospital admissions after prostate biopsy in a European randomized trial. Eur Urol. 2012; 61:1110–4. [PubMed: 22244150]
54. Loeb S, Vellekoop A, Ahmed HU, et al. Systematic review of complications of prostate biopsy. Eur Urol. 2013; 64:876–92. [PubMed: 23787356]
55. Alibhai SM, Leach M, Tomlinson G, et al. 30-day mortality and major complications after radical prostatectomy: influence of age and comorbidity. J Natl Cancer Inst. 2005; 97:1525–32. [PubMed: 16234566]
56. US Preventive Services Task Force Draft Evidence Review for Prostate Cancer Screening. https://www.uspreventiveservicestaskforce.org/Page/Document/draft-evidence-review/prostate-cancer-screening1

57. Bremner KE, Chong CA, Tomlinson G, Alibhai SM, Krahn MD. A review and meta-analysis of prostate cancer utilities. Med Decis Making. 2007; 27:288–98. [PubMed: 17502448]

Take Home Message

Active surveillance extends life more than watchful waiting, particularly for men with higher-risk features, but this is partly offset by decrements in quality of life from delayed treatment. Trade-offs about the intensity of surveillance should be discussed with patients.

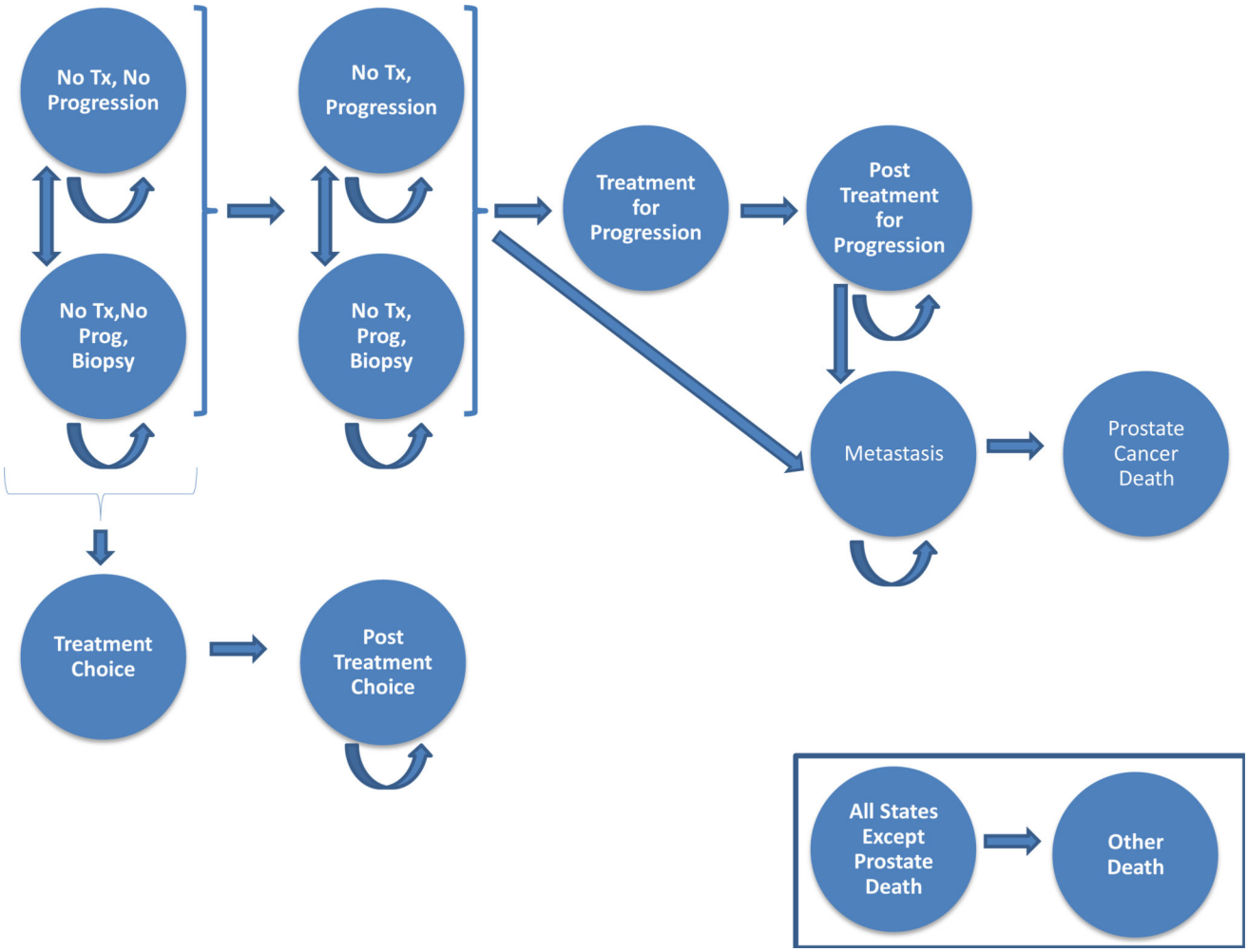
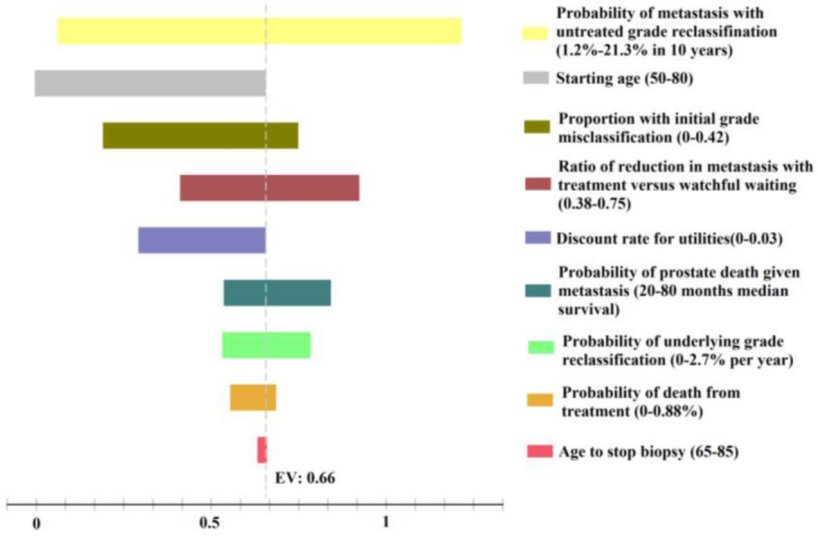


Fig. 1.

Schematic diagram of the state-transition Markov model for men undergoing conservative management of prostate cancer showing all the possible states that men in the model can be in and all the possible transitions between states. At the start, men have been diagnosed with PCa and have chosen conservative management. Some were classified accurately with Gleason 6 (grade group 1), while others were misclassified and have undetected higher-grade disease. During each model cycle, individuals can remain on conservative management, undergo treatment for reclassification (then into a post-treatment state), develop metastases, or die. PCa = prostate cancer; Prog = progression; Tx = treatment. Biopsy, treatment and post-treatment states are silent during watchful waiting. In the efficacy analysis shown in this paper, patients only undergo treatment for evidence of reclassification.

(a) Tornado diagram showing one-way sensitivity analyses for the outcome life-years



(b) Tornado diagram showing one-way sensivity analyses for the outcome quality-adjusted life-years

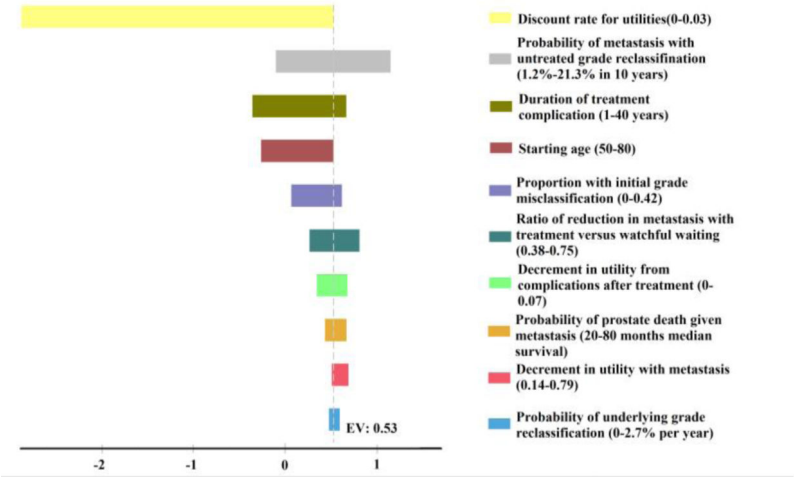
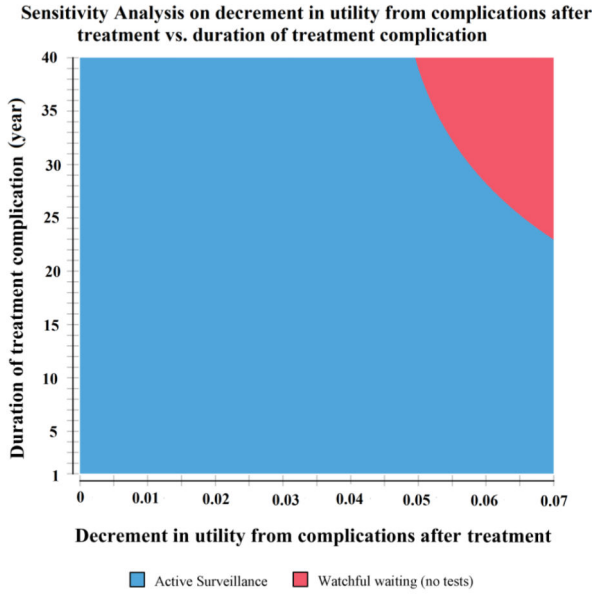


Fig. 2.

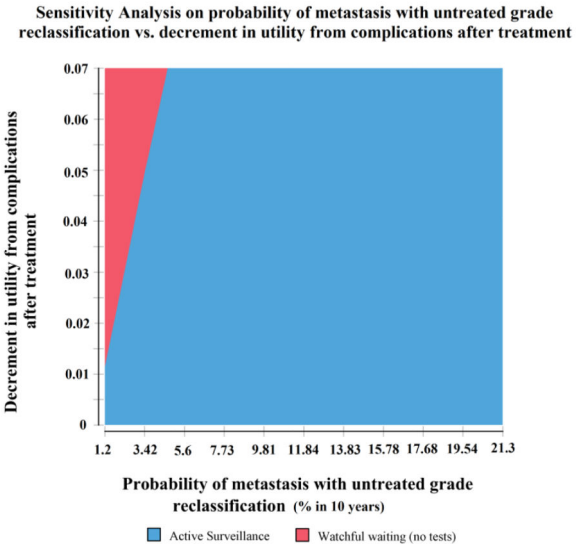
Tornado diagram showing a series of one-way sensitivity analyses of key variables for the outcome of (A) LYs and (B) QALYs comparing AS (Johns Hopkins) with WW. The tornado diagram for incremental LYs (or QALYs) shows how the difference in LYs (or QALYs) between AS and WW changes when the value of a parameter varies. The *X*-axis shows the difference in LYs (or QALYs) between AS and WW. The dotted line shows the difference in LYs (or QALYs) for the base case analysis, where AS has 0.66 more LYs (or 0.53 more QALYs) than WW. Each bar shows how much the difference in LYs (or QALYs) changes when we change a specific parameter within its range. If a bar crosses “0” in *X*-axis, it

means that AS has less LYs (or QALYs) than WW and therefore the decision is reversed. AS = active surveillance; LY = life year; QALY = quality-adjusted life year; WW = watchful waiting.

(A)



(B)



(C)

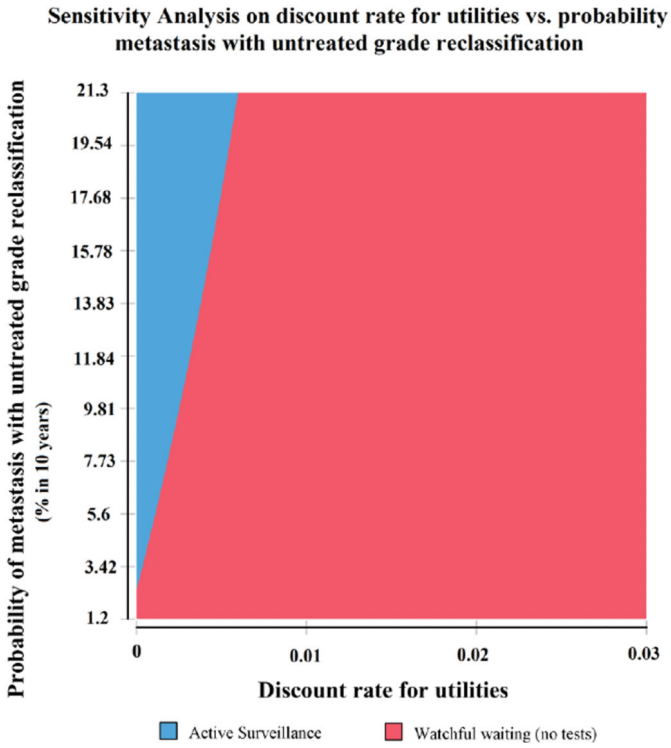


Fig. 3.

Two-way sensitivity analyses for the (A) decrement in utility from treatment complications and the duration of treatment-related complications, (B) probability of metastasis for untreated grade reclassification and decrement in utility from treatment-related complications, and (C) discount rate and probability of metastasis with untreated grade reclassification. Active surveillance (Hopkins) is preferred with a shorter duration and less utility decrement from treatment complications, and with an increasing probability of metastasis for untreated grade reclassification, whereas watchful waiting has more QALYs with a large decrement in utility and long duration of treatment-related complications, and with a higher discount rate. QALY = quality-adjusted life year.

Table 1

Parameters of the Markov model comparing watchful waiting and active surveillance

Variable	Point estimate	Range for sensitivity analysis[a]
Epidemiologic variables		
Proportion with initial grade misclassification [39]	35% low risk	0–42%
	31% very low risk	
Probability of grade reclassification [40,41]	Low risk: 1.2%/yr	0–2.7%/yr
	Very low risk: 1%/yr	
Probability of metastasis with untreated grade reclassification [42–45]	MFS 99% at 5 yr, 91% at 10 yr, 82% at 15 yr, then stabilizes	1.2–21.3% at 10 yr
Relative risk of metastasis with treatment versus watchful waiting [8,26,42]	0.57	0.38–0.75
Probability of PCa death given metastasis [46–49]	Median overall survival 60 mo, 85% PCa death	20–80 mo
Test performance variables		
PSA sensitivity [25]	49.5%	40.2–58.8%
PSA specificity [25]	50.8%	44.2–78.7%
MRI sensitivity [50]	69%	44–86%
MRI specificity [50]	78%	53–91%
Biopsy sensitivity with normal MRI [51,52]	53%	43–63%
Increase in biopsy sensitivity with an abnormal MRI [51]	32%	23–38%
Biopsy specificity	1	Fixed at 1 (assumption)
Complications variables		
Probability of infection after biopsy[53,54]	4.0%	0–6.3%
Probability of death from treatment [6,8,26,55,56]	0.2%	0–0.88%
Utilities [b]		
Utility for no treatment [16]	0.97	0.5–1
Decrement in utility for patients having complication after biopsy [16][c]	0.07	0.06–0.43
Utility during treatment [16]	0.67	0.65–0.90
Decrement in utility from complications after treatment [6,16,57] [d]	0.02	0–0.29
Duration of utility decrement from complications after treatment [16]	10 yr	1–40 yr
Decrement in utility with metastasis [6,16,57]	0.21	0.10–0.50
Discount rate	0	0–0.03

MFS = metastasis-free survival; MRI = magnetic resonance imaging; PCa = prostate cancer; PSA = prostate-specific antigen.

[a] The range for sensitivity analysis was drawn from the literature.

[b] Utility decrements were used to preserve the rank order of utilities for different states. For example, the utility for “no treatment, biopsy” is defined by subtracting the decrement of utility for the “no treatment, biopsy” state from the utility of the “no treatment” state. By setting the upper bound of the decrement to be less than the utility of the “no treatment” state, we can assure that the utility for the “no treatment, biopsy” state is always lower than that for the “no treatment” state.

[c] The decrement in utility for biopsy complications was applied for 1 mo.

[d] A decrement of 0.11 was applied to men undergoing treatment at age >70 to account for more frequent complications in this age group.

Table 2

Comparisons of remaining life expectancy and quality-adjusted life expectancy between active surveillance using different protocols for men with low-risk prostate cancer, compared with watchful waiting in the cohorts aged ≥40 yr, ≥50 yr (base case), ≥65 yr, ≥70 yr, and ≥75 yr

Strategy	Remaining life expectancy (LY)	Incremental LY	Quality-adjusted life expectancy (QALY)	Incremental QALY
Cohort aged ≥40 yr				
Watchful waiting	42.94	–	41.47	–
AS—Hopkins	43.96	+1.03	42.36	+0.89
AS—PRIAS	43.81	+0.88	42.20	+0.73
AS—MRI based	43.96	+1.03	42.36	+0.90
AS—5 yr	43.58	+0.64	41.95	+0.49
Cohort aged ≥50 yr				
Watchful waiting	34.55	–	33.36	–
AS—Hopkins	35.21	+0.66	33.89	+0.53
AS—PRIAS	35.12	+0.57	33.79	+0.44
AS—MRI based	35.20	+0.65	33.89	+0.53
AS—5 yr	34.99	+0.44	33.63	+0.27
Cohort aged ≥65 yr				
Watchful waiting	22.60	–	21.80	
AS—Hopkins	22.83	+0.24	21.70	−0.10
AS—PRIAS	22.81	+0.22	21.70	−0.10
AS—MRI based	22.83	+0.24	21.71	−0.10
AS—5 yr	22.78	+0.19	21.67	−0.13
Cohort aged ≥70 yr				
Watchful waiting	18.87	–	18.21	–
AS—Hopkins	19.02	+0.14	17.89	−0.31
AS—PRIAS	19.00	+0.13	17.93	−0.28
AS—MRI based	19.01	+0.14	17.89	−0.32
AS—5 yr	18.99	+0.12	18.00	−0.20
Cohort aged ≥75 yr				
Watchful waiting	15.35	–	14.81	–
AS—Hopkins	15.42	+0.07	14.48	−0.33
AS—PRIAS	15.41	+0.06	14.52	−0.29
AS—MRI based	15.42	+0.07	14.47	−0.34
AS—5 yr	15.41	+0.06	14.63	−0.18

AS = active surveillance; LY = life years; QALY = quality-adjusted life years; MRI = magnetic resonance imaging; PRIAS = Prostate Cancer Research International Active Surveillance.