Systematic Review and Meta-analysis of Factors Determining Change to Radical Treatment in Active Surveillance for Localized Prostate Cancer

Andrew J. Simpkin, Kate Tilling, Richard M. Martin, J. Athene Lane, Freddie C. Hamdy, Lars Holmberg, David E. Neal, Chris Metcalfe, Jenny L. Donovan

School of Social and Community Medicine, University of Bristol, Bristol, UK; NIHR Bristol Nutrition Biomedical Research Unit, University of Bristol, Bristol, UK; Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK; Guy's Hospital, London, UK; Cancer Research UK, Cambridge Research Institute, Cambridge, UK

Abstract

Context: Many men with clinically localized prostate cancer are being monitored as part of active surveillance (AS) programs, but little is known about reasons for receiving radical treatment.

Objectives: A systematic review of the evidence about AS was undertaken, with a meta-analysis to identify predictors of radical treatment.

Evidence acquisition: A comprehensive search of the Embase, MEDLINE and Web of Knowledge databases to March 2014 was performed. Studies reporting on men with localized prostate cancer followed by AS or monitoring were included. AS was defined where objective eligibility criteria, management strategies, and triggers for clinical review or radical treatment were reported.

Evidence synthesis: The 26 AS cohorts included 7627 men, with a median follow-up of 3.5 yr (range of medians 1.5–7.5 yr). The cohorts had a wide range of inclusion criteria, monitoring protocols, and triggers for radical treatment. There were eight prostate cancer deaths and five cases of metastases in 24 981 person-years of follow-up. Each year, 8.8% of men (95% confidence interval 6.7–11.0%) received radical treatment, most commonly because of biopsy findings, prostate-specific antigen triggers, or patient choice driven by anxiety. Studies in which most men changed treatment were those including only low-risk Gleason score 6 disease and scheduled rebiopsies.

Conclusions: The wide variety of AS protocols and lack of robust evidence make firm conclusions difficult. Currently, patients and clinicians have to make judgments about the balance of risks and benefits in AS protocols. The publication of robust evidence from randomized trials and longer-term follow-up of cohorts is urgently required.

Patient summary: We reviewed 26 studies of men on active surveillance for prostate cancer. There was evidence that studies including men with the lowest risk disease and scheduled rebiopsy had higher rates of radical treatment.
1. Introduction

Active surveillance (AS) is increasingly used as an alternative to immediate radical intervention for men with clinically localized prostate cancer who are at low risk of progressing to life-threatening disease [1]. Radical treatment comes with a fairly high risk of harm [2], so there is strong motivation to radically intervene only in cases with high risk of progression. AS involves regular follow-up with prostate-specific antigen (PSA) testing, digital rectal examination (DRE), review of symptoms, and/or repeat biopsy. Surveillance also requires predefined triggers for clinical review so that radical treatment can be initiated where possible for those with progressing disease. There is no formal evidence on whether AS is a safe management option for men with clinically localized prostate cancer. Two randomized trials have evaluated the effectiveness of a passive strategy called watchful waiting (ie, palliative treatment once symptoms appear). The Prostate Cancer Intervention versus Observation Trial (PIVOT) recently found no difference between watchful waiting and radical prostatectomy for either all-cause or prostate cancer mortality after at least 10 yr of follow-up [3]. The PIVOT cohort was relatively elderly (mean age 67 yr at enrolment) and the majority had screen-detected low-risk disease. The SPCG-4 study [4,5] found that radical prostatectomy reduced prostate-specific mortality compared to watchful waiting after a median of 13.2 yr among men who had been diagnosed clinically.

Previous systematic reviews [6–8] have found little consensus on eligibility criteria, protocols for surveillance, or triggers for recommending radical intervention. These various strategies for AS have led to widely different rates of change to radical treatment between AS studies [6–8]. We undertook a systematic review and meta-analysis to investigate rates of change to radical treatment and the key AS strategy factors influencing change of treatment.

2. Evidence acquisition

2.1. Search strategy

We conducted a systematic search of the MEDLINE, Embase, and Web of Science online databases from October 2004 (the end date of our previous systematic review [6]) to April 2013. A forward citation search of the five studies [1,9–12] included by Martin et al [6] was also performed using the Web of Science database. An update using the same search strategy was carried out by A.J.S. alone in March 2014. Further information about our evidence acquisition and synthesis can be found in the Supplementary material online.

2.2. Inclusion criteria

We included studies involving men with T1–T2 clinically localized prostate cancer that was initially managed conservatively, and in which predefined clinical, pathological, or biochemical criteria for clinical review were outlined. Studies involving recurrence after radical prostatectomy or radiotherapy (ie, not initially managed conservatively) were excluded. We excluded any reviews, editorial comments, background papers, and studies involving different treatments and diseases of the prostate.

2.3. Data extraction and synthesis

Eligibility criteria, surveillance protocols, sample size, age, PSA, follow-up times, treatment change triggers, treatment change rates, metastases, prostate cancer-specific mortality, and reasons for changing treatment were extracted manually from each paper by A.J.S. and checked by one of C.M., K.T., or R.M.M. Authors of publications found in our search were contacted to provide further data where necessary and to check that data extraction was correct.

2.4. Study outcomes

Treatment change rates are considered here as key short-term outcome measures for AS, and the occurrence of metastases and/or prostate cancer death are longer-term outcomes. To account for both sample size and duration of follow-up in the cohorts, person-years were used in calculating the rate of change to radical treatment. Person-years were estimated as the median follow-up time multiplied by the sample size for each study; total person-years were the sum of these across the studies included.

2.5. Meta-analysis methods

A meta-analysis was conducted to estimate the rate of change to radical treatment per person-year. This rate was calculated for each study as

\[
\text{rate of change to radical treatment per person-year} = \frac{\text{radical treatment events}}{\text{sample size} \times \text{median follow-up time}}.
\]

Heterogeneity between studies was measured using the I^2 statistic [13]: a higher I^2 value indicates higher between-study heterogeneity. Meta-regression was performed to examine the associations of study characteristics with rates of change to radical treatment. The study characteristics considered were (1) year of first recruitment; (2) whether the study restricted participation to those with Gleason grade 3 + 3 or less; (3) whether the study limited inclusion to men with PSA ≤ 10 ng/ml; (4) the number of scheduled PSA tests in the first 3 yr; (5) whether the study protocol included scheduled rebiopsy; and (6) whether PSA or PSA kinetic measures, such as PSA doubling time (PSADT) or PSA velocity (PSAV), were used to recommend clinical review or radical treatment. Together with univariate analysis, two multivariable meta-regressions were performed, grouping eligibility variables (1)–(3) and monitoring procedure variables (4)–(6).

Several studies had conducted within-cohort analyses to relate patient characteristics to the change to radical treatment. For cases in which more than two studies reported estimates of the association of a patient characteristic with change to radical treatment, we carried out a meta-analysis of these estimates. We also used meta-analysis to examine the reasons for changing to radical
treatment. Reasons were given within studies and the proportions of each reason were combined to obtain an overall average estimate.

2.6. Bias in reporting

To examine whether the reviewed studies were subject to a small-study effect (whereby smaller studies give different estimates of the risk of radical treatment) that might be due to publication bias (smaller studies are more likely to be published if their results are extreme) we used plots of percentage change to radical treatment against sample size.

3. Evidence synthesis

Figure 1 shows a PRISMA (Preferred Reporting Items of Systematic Reviews and Meta-analyses) flow chart of the screening and selection results for this review. We included 26 AS studies comprising 7627 men. The median age at entry ranged from 63 to 70.5 yr, and the median follow-up time between the cohorts was 3.5 yr (range 1.5–7.5 yr). Overall, eight prostate cancer deaths were reported [14–16] among the 7627 men and 24,981 person-years of follow-up, with six further men developing metastases [10,15,23,27].

3.1. Eligibility criteria, surveillance protocols, and triggers for change to radical treatment

Studies used explicit eligibility criteria involving Gleason grade (23 studies) [12,14–35], T stage (22) [9,10,15–26, 28,30–36], or PSA or PSA density (19) [14–20,22,23,25,26, 28–35], with varying thresholds (Table 1). All studies used regular PSA testing. Rebiopsy was used in 23 of the 26 studies [10,12,14–29,32–36], mostly with regular rebiopsy intervals, although in three studies rebiopsy was only performed in cases with worsening PSA or DRE results [10,15,36]. The frequency of PSA testing, DRE, and rebiopsy during surveillance varied considerably across the studies. An increase in Gleason grade [12,14–29,32–36] was the most common trigger for recommendation of a change to radical treatment, followed by other pathological findings [12,14,16–24, 26,27,29,32–36], PSA or PSA kinetics [9,10,12,14,16,17, 19,23,26,28–30,32,34,35], and T stage [14,15,17,19,23,26].

3.2. Radical treatment rates

The average rate of change to radical treatment was 88 per 1000 person-years (95% confidence interval [CI] 67, 110). This corresponds to an estimated 88 out of 250 men on surveillance receiving radical treatment during 4 yr of follow-up (Fig. 2). There was evidence ($I^2 = 96%; p < 0.0005$) that the rate of change differed substantially between cohorts. The minimum rate was 11 per 1000 person-years [10] and the maximum was 218 per 1000 person-years [18].

3.3. Study characteristics associated with the rate of change to radical treatment

The rate of change was greater in studies with recruitment periods that started in more recent years; for each year increase in start date, an extra four men changed treatment per 1000 person-years (95% CI 1, 6; $p = 0.008$; Table 2). Studies admitting men with higher Gleason score had an average decrease of 56 treatment changes per 1000 person-years (95% CI 20, 93; $p = 0.004$; Table 2) compared to studies only including men with Gleason 6 or lower (Fig. 3). In studies with scheduled rebiopsy, an average of 53 more men changed treatment per 1000 person-years (95% CI 9, 98; $p = 0.021$; Table 2) compared to studies that biopsied only on worsening results for PSA or DRE (Fig. 3). In multivariate
<table>
<thead>
<tr>
<th>Setting</th>
<th>Sample size (recruited)</th>
<th>Eligibility</th>
<th>Study design</th>
<th>Surveillance protocol</th>
<th>Trigger for clinical review and treatment change criteria</th>
<th>Median follow-up, yr (range)</th>
<th>Patients changing to radical treatment, % (median time, yr)</th>
<th>Reasons for changing treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taichung Veterans General Hospital, Taiwan [10]</td>
<td>52 (1983–1996)</td>
<td>–</td>
<td>1a</td>
<td>TURP for benign hyperplasia</td>
<td>PSA &gt;4 ng/ml or PSAv &gt; 1 ng/ml/yr leads to rebiopsy where upgrade can lead to treatment</td>
<td>7 (0.5–15)</td>
<td>8 (2.5)</td>
<td>4 abnormal DRE and PSA elevation</td>
</tr>
<tr>
<td>Memorial Sloan Kettering Cancer Centre, New York, USA [12]</td>
<td>88 (1984–2001)</td>
<td>–</td>
<td>4 + 3</td>
<td>Restaging biopsy after diagnosis, then due to DRE/PSA</td>
<td>Any of predominant GS 4 pattern, &gt;2 positive cores, &gt;50% cancer/core, ≥ T2b</td>
<td>6.3 (1.7–14.1)</td>
<td>16 (3.7)</td>
<td>9 ≥ T2b 13 &gt;50% cancer/core 5 Gleason 22 &gt;3 positive cores 2 patient choice</td>
</tr>
<tr>
<td>Hospital Universitario Miguel Servet, Zaragoza, Spain [24]</td>
<td>16 (1986–1999)</td>
<td>–</td>
<td>3 + 3</td>
<td>At 3 mo, then every 6 mo</td>
<td>Any of increase in GS, increase in no. of positive cores, onset of urinary symptoms, change in DRE, patient request</td>
<td>7.5</td>
<td>0</td>
<td>Not specified</td>
</tr>
<tr>
<td>McGill University, Canada [27]</td>
<td>186 (1987–2006)</td>
<td>–</td>
<td>3 + 3</td>
<td>Within 2 mo, then every 6 mo for 2 yr, then every 12 mo</td>
<td>Doubling of PSA from baseline leads to rebiopsy where upgrade can lead to treatment</td>
<td>5.1</td>
<td>19</td>
<td>Not specified</td>
</tr>
<tr>
<td>Cochin Hospital, Paris, France [36]</td>
<td>144 (1988–2005)</td>
<td>–</td>
<td>1a</td>
<td>Not routinely taken</td>
<td>Abnormal DRE or 3 consecutive PSA increases with total increase of 5 ng/ml</td>
<td>Mean 1.9 (0.5, 5.1)</td>
<td>19 (1)</td>
<td>5 PSA increases</td>
</tr>
<tr>
<td>University of Connecticut, USA [30]</td>
<td>40 (1990–2006)</td>
<td>≤10</td>
<td>3 + 3</td>
<td>Every 6 mo or due to DRE/PSA</td>
<td>Abnormal DRE or progressive PSA elevation</td>
<td>4 (1–14)</td>
<td>23 (2.75)</td>
<td>Not specified</td>
</tr>
<tr>
<td>UNC, Chapel Hill, USA [9]</td>
<td>27 (1991–1996)</td>
<td>–</td>
<td>1c</td>
<td>At 3 mo, then every 6 mo</td>
<td>Abnormal DRE or 3 consecutive PSA increases with total increase of 5 ng/ml</td>
<td>3.4</td>
<td>29 (2.5)</td>
<td>10 based on rate of PSA rise or subjective decision by patient and clinician</td>
</tr>
<tr>
<td>ERSPC, Rotterdam, Holland [31]</td>
<td>278 (1993–1999)</td>
<td>≤15</td>
<td>4 + 3</td>
<td>Chart reviews every 6 mo</td>
<td>Either rate of PSA increase or subjective advice</td>
<td>3.5 (0.1–9.7)</td>
<td>14</td>
<td>Not specified</td>
</tr>
<tr>
<td>Royal Marsden NHS Trust, UK [30]</td>
<td>80 (1993–2002)</td>
<td>≤20</td>
<td>4 + 3</td>
<td>Every 3–6 mo, then every 6 mo after 2 yr</td>
<td>Not routinely taken</td>
<td>3.4</td>
<td>29 (2.5)</td>
<td>10 based on rate of PSA rise or subjective decision by patient and clinician</td>
</tr>
<tr>
<td>Institution</td>
<td>PSA Cut-off</td>
<td>PSA Density</td>
<td>GS Cut-off</td>
<td>Surveillance Frequency</td>
<td>PSADT Cut-off</td>
<td>T-stage Cut-off</td>
<td>GS Increase Cut-off</td>
<td>Imaging Frequency</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td>------------------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Memorial Sloan Kettering Cancer Center, New York, USA [17]</td>
<td>&lt;10 ng/ml</td>
<td>2a</td>
<td>3 + 3 positive cores (out of at least 10) &lt;50% cancer in any core, confirmatory biopsy before starting surveillance</td>
<td>Every 6 mo</td>
<td>Every 2-3 yr or due to DRE/PSA</td>
<td>Within 12-18 mo, then 50% tumor/core, &gt;T2a</td>
<td>1.8 for untreated; 11% ≥5 yr</td>
<td>23 GS ≥7</td>
</tr>
<tr>
<td>University of Miami, USA [22]</td>
<td>≤10 ng/ml</td>
<td>2a</td>
<td>≤2 cores ≤20% cancer in any core</td>
<td>Every 3–4 mo for 2 yr, then every 6 mo</td>
<td>Every 3–4 mo for 2 yr, then every 6 mo</td>
<td>Within 1 yr, then every 1–2 yr</td>
<td>Any of GS ≥7, increase in no. of positive cores, increase in % cancer/core or personal choice</td>
<td>3.3 (1–17.3)</td>
</tr>
<tr>
<td>University of Toronto, Canada [14]</td>
<td>≤10 ng/ml</td>
<td>0.15 ng/ml/cm²</td>
<td>≤3 positive cores, &lt;50% involvement in any core</td>
<td>Every 3–6 mo</td>
<td>Every 3–6 mo</td>
<td>Due to signs of PSA or T-stage progression</td>
<td>Established PSA, stage, or grade progression</td>
<td>6.8</td>
</tr>
<tr>
<td>ERSPC, Gothenburg, Sweden [15]</td>
<td>&lt;10 ng/ml</td>
<td>1 (LRG)</td>
<td>Every LRG defined as T1c, GS ≤6, PSAd &lt;0.15 ng/ml/cm², &lt;3 positive cores, ≤50% involvement in any core</td>
<td>Every 3–6 mo</td>
<td>Every 3–6 mo</td>
<td>Change in no. of positive cores, increase in % cancer/core</td>
<td>3.9 (1–15.2)</td>
<td>27</td>
</tr>
<tr>
<td>UC, San Francisco, USA [28]</td>
<td>≤0.15 ng/ml/cm²</td>
<td>3 + 3 (LRG)</td>
<td>LRG: GS &lt; 4 + 3 if already or to 3 + 4 otherwise</td>
<td>Every 3 mo</td>
<td>Every 12–24 mo</td>
<td>Upgrade to PSADT &lt; 2 yr, GS ≥4 + 3 or to 3 + 4 otherwise</td>
<td>30% of LRG (4) 35% of IRG (4) 73 reclassified on repeat biopsy</td>
<td>14</td>
</tr>
<tr>
<td>Johns Hopkins University, USA [19]</td>
<td>PSAd &lt; 0.15 ng/ml/cm²</td>
<td>3 + 3</td>
<td>≤2 positive cores ≤50% cancer in any core</td>
<td>Every 6 mo</td>
<td>Every 6 mo</td>
<td>Any of PSAd ≥0.15 ng/ml/cm², GS ≥7, &gt;2 positive cores, &gt;50% tumor/core</td>
<td>2.7 (0.01–15)</td>
<td>33</td>
</tr>
<tr>
<td>Harvard University, USA [21]</td>
<td>–</td>
<td>3 + 3</td>
<td>≤2 positive cores ≤50% cancer in any core</td>
<td>Every 6 mo</td>
<td>Every 6 mo</td>
<td>Any of GS ≥7, ≥3 positive cores, &gt;50% tumor/core</td>
<td>2.4</td>
<td>27</td>
</tr>
<tr>
<td>Shikoku Cancer Centre, Matsuyama, Japan [35]</td>
<td>≤10 ng/ml</td>
<td>3 + 3</td>
<td>≤2 positive cores ≤50% cancer in any core</td>
<td>Every 1–3 mo</td>
<td>Every 6 mo</td>
<td>At 12 mo, then every 1–2 yr</td>
<td>Any of GS ≥7, ≥3 positive cores, &gt;50% tumor/core</td>
<td>2.9</td>
</tr>
<tr>
<td>Multi-institutional, Kagawa, Japan based [26]</td>
<td>≤20 ng/ml</td>
<td>3 + 3</td>
<td>≤2 positive cores ≤50% cancer in any core age &lt;80 yr</td>
<td>Every 2 mo for 6 mo, then every 3 mo</td>
<td>At 12 mo</td>
<td>PSADT ≥2 yr or pathological upgrade on rebiopsy</td>
<td>4.5</td>
<td>47</td>
</tr>
<tr>
<td>Setting</td>
<td>Sample size (recruited)</td>
<td>Eligibility</td>
<td>Study design</td>
<td>Surveillance protocol</td>
<td>Trigger for clinical review and treatment change criteria</td>
<td>Median follow-up yr (range)</td>
<td>Patients changing to radical treatment, % (median time, yr)</td>
<td>Reasons for changing treatment</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-----------------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Royal Marsden NHS Trust, UK [16]</td>
<td>471 (2002–2011)</td>
<td>&lt;15</td>
<td>3 + 4</td>
<td>≤50% positive cores</td>
<td>Any of PSAv &gt;1 ng/ml/yr, GS ≥ 4 + 3, &gt;50% tumor/core</td>
<td>3.7 (1.6)</td>
<td>31 (5.4)</td>
<td>18 history alone 56 PSAv alone 23 both history and PSAv 40 patient choice</td>
</tr>
<tr>
<td>Southern Health, Melbourne, Australia [29]</td>
<td>154 (2003–2010)</td>
<td>&lt;10</td>
<td>3 + 4</td>
<td>≤2 positive cores</td>
<td>PSADT &lt;3 yr leads to re-biopsy, any Gleason or mm cancer/core increase</td>
<td>1.9 (0.1-16.6)</td>
<td>17 (2.4)</td>
<td>26 reclassified on repeat biopsy 1 patient choice</td>
</tr>
<tr>
<td>Hospital Universitario Fundacion de Alcorcon, Madrid, Spain [34]</td>
<td>144 (2004–2012)</td>
<td>&lt;10</td>
<td>3 + 3</td>
<td>≤33% positive cores</td>
<td>Either pathological progression or PSADT &lt;3 yr</td>
<td>3.2 (2.1)</td>
<td>24</td>
<td>25 Gleason/cores 5 patient choice 4 PSA</td>
</tr>
<tr>
<td>Cleveland, Detroit, USA and Guadalupe [25]</td>
<td>139 (2005–2012)</td>
<td>≤10</td>
<td>3 + 3</td>
<td>≤3 positive cores</td>
<td>Any of GS ≥7, &gt;33% positive cores, &gt;50% cancer in any core</td>
<td>3.3 (2.1)</td>
<td>26</td>
<td>14 patient choice 10 Gleason 8 no. of cores 9% cancer/core 1 PSA 3 unknown</td>
</tr>
<tr>
<td>REDEEM Study control arm, multi-institutional, North America [33]</td>
<td>155 (2006–2007)</td>
<td>≤11</td>
<td>3 + 3</td>
<td>≤3 positive cores</td>
<td>Any of GS ≥7, &gt;3 positive cores, &gt;50% tumor/core</td>
<td>3 (3.7)</td>
<td>37.5</td>
<td>7 Gleason 30 cores 14 Gleason and cores</td>
</tr>
<tr>
<td>Frimley Park Hospital, UK [18]</td>
<td>101 (2006–2010)</td>
<td>≤15</td>
<td>3 + 3</td>
<td>≤50% positive cores</td>
<td>Any of upgrade on TT, increase in GS, &gt;50% positive cores, &gt;10 mm cancer/core</td>
<td>1.5 (1–2.3)</td>
<td>33</td>
<td>25/34 reclassified on TTB chose treatment 10 patient choice</td>
</tr>
<tr>
<td>PRIAS (International), Rotterdam based, Holland [23]</td>
<td>2494 (2006–2012)</td>
<td>≤10</td>
<td>3 + 3</td>
<td>≤2 positive cores</td>
<td>Any of PSADT &lt;3 yr, stage – T2</td>
<td>1.6 (IQ 1.0–2.8)</td>
<td>21</td>
<td>387 on biopsy or PSADT 47 anxiety 93 other</td>
</tr>
<tr>
<td>Beaumont Hospital, Dublin, Ireland [32]</td>
<td>80 (2008–2012)</td>
<td>≤10</td>
<td>3 + 3</td>
<td>≤2 positive cores</td>
<td>Any of PSADT &lt;3 yr, stage – T2</td>
<td>3.1 (range 0.2–4.5)</td>
<td>42.5</td>
<td>23 reclassified on repeat biopsy 6 PSA 4 MRI 1 DRE</td>
</tr>
</tbody>
</table>

DRE = Digital rectal exam; PSA = prostate specific antigen, PSADT = PSA doubling time; PSAd = PSA density; CAPRA score = a score used by the UC San Francisco cohort to rate prostate cancer, includes PSA, age, T stage, Gleason score, and percentage of biopsy cores involved with cancer; GS = Gleason score; LRG = low-risk group; IRG = intermediate-risk group; MRI = magnetic resonance imaging; TTB = transperineal template biopsy; TURP = transurethral radical prostatectomy; IQR = interquartile range; PSAv = PSA velocity.

* 37.5% had pathological progression at 3 yr.
meta-regression with mutual adjustment for all covariates, the results were slightly attenuated, although some evidence for associations with Gleason grade > 3 + 3 (53 fewer changes per 1000 person-years, 95% CI 3, 102; \( p = 0.038 \)) and regular biopsy (45 more changes per 1000 person-years, 95% CI –4, 94; \( p = 0.073 \)) still remained.

### 3.4. Individual factors associated with change to radical treatment

Seven studies examined patient characteristics in relation to the time to a change to radical treatment [16,17,21,27,36] or the rate of change to radical treatment [14,23]. The predictors that were common to more than two of these seven studies were baseline PSA, Gleason grade (\(< 7 \text{ vs } \geq 7\)) and T stage (\(< 2a \text{ vs } \geq 2a\)). Forest plots displaying the pooled effect estimates are presented in the Supplementary material. The effect of PSA was heterogeneous among four studies [16,17,23,36] (\( I^2 = 92\%, p < 0.0005 \)) and the pooled hazard ratio (HR) for change to radical treatment was 1.03 per ng/ml higher PSA (95% CI 1.01, 1.06; \( p = 0.02 \)); baseline Gleason grade > 3 + 3 increased the chance of changing to radical treatment in four of the studies [14,16,27,36] (pooled HR 1.98, 95% CI 1.54, 2.55; \( p < 0.0005 \); Table 3). Baseline T stage \( \geq 2a \) was associated with increased treatment change in two studies [14,16], and combining these with three inconclusive studies [17,23,27]...

---

**Table 2 – Results from meta-regression for associations between study design characteristics and rate of treatment change**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Unadjusted change in rate per 1000 person-years (95% CI)</th>
<th>p value</th>
<th>Adjusted change in rate per 1000 person-years (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruitment year (mean change in rate per year from 1983)</td>
<td>4 (1, 6)</td>
<td>0.008</td>
<td>1 (–2, 4)</td>
<td>0.55</td>
</tr>
<tr>
<td>Gleason (difference in rate between cohorts with GS ( \leq 3 + 3 ) and cohorts with GS ( \leq 4 + 3 ); a positive value implies a higher rate in the former)</td>
<td>56 (20, 93)</td>
<td>0.004</td>
<td>53 (3, 102)</td>
<td>0.038</td>
</tr>
<tr>
<td>PSA (difference in rate between cohorts with PSA ( \leq 10 \text{ ng/ml} ) and cohorts with PSA &gt;10 ng/ml; a positive value implies a higher rate in the former)</td>
<td>39 (–3, 82)</td>
<td>0.072</td>
<td>4 (–43, 50)</td>
<td>0.87</td>
</tr>
<tr>
<td>PSA frequency in first 3 yr (mean change per extra PSA test taken)</td>
<td>4 (–1, 9)</td>
<td>0.077</td>
<td>3 (–2, 8)</td>
<td>0.26</td>
</tr>
<tr>
<td>Regular biopsy (mean increase in rate using a scheduled biopsy)</td>
<td>53 (9, 98)</td>
<td>0.021</td>
<td>45 (–4, 94)</td>
<td>0.073</td>
</tr>
<tr>
<td>PSA progression (mean increase in rate using PSA as a trigger)</td>
<td>–3 (–52, 47)</td>
<td>0.91</td>
<td>–0.4 (–46, 46)</td>
<td>0.99</td>
</tr>
</tbody>
</table>

CI = confidence interval; GS = Gleason score; PSA = prostate-specific antigen.

---

**Fig. 2 – Forest plot of the proportion of men changing to radical treatment per person years, sorted by initial year of recruitment (in parentheses after study setting). References for the study settings are provided in Table 1.**

**Fig. 3 – Treatment change rate per 1000 person-years against follow-up time, grouped by (A) Gleason entry criteria and (B) rebiopsy procedure.**

**Fig. 2** - Forest plot of the proportion of men changing to radical treatment per person years, sorted by initial year of recruitment (in parentheses after study setting). References for the study settings are provided in Table 1.

**Fig. 3** - Treatment change rate per 1000 person-years against follow-up time, grouped by (A) Gleason entry criteria and (B) rebiopsy procedure.
References for the study settings are provided in Table 1. After the study setting, some confidence intervals have negative values because of small sample sizes and consequent invalid normal approximations.

Gleason score increase, (B), PSA increase, (C), other biopsy reclassification, and (D) choice or anxiety. The year of initial recruitment is in parentheses.

Table 3 – Odds or hazard ratios from common covariates in models for changing to radical treatment, with 95% confidence intervals in parentheses

<table>
<thead>
<tr>
<th>Setting</th>
<th>Model</th>
<th>Predictors of progression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Per unit increase in baseline PSA (ng/ml)</td>
</tr>
<tr>
<td>McGill University, Canada [27]</td>
<td>TRT</td>
<td>2.70 (1.40, 5.50)</td>
</tr>
<tr>
<td>Cochin Hospital, Paris [36]</td>
<td>TRT</td>
<td>1.07 (1.03, 1.11)</td>
</tr>
<tr>
<td>Memorial Sloan Kettering Cancer Center, New York [17]</td>
<td>TRT</td>
<td>1.31 (1.17, 1.47)</td>
</tr>
<tr>
<td>University of Toronto, Canada [14]</td>
<td>RT</td>
<td>–</td>
</tr>
<tr>
<td>Harvard University, USA [21]</td>
<td>TRT</td>
<td>–</td>
</tr>
<tr>
<td>Royal Marsden NHS Trust, UK [16]</td>
<td>TRT</td>
<td>1.0 (1.0, 1.1)</td>
</tr>
<tr>
<td>PRIAS, International [23]</td>
<td>RT</td>
<td>0.90 (0.84, 0.96)</td>
</tr>
</tbody>
</table>

TRT = time to radical treatment; RT = radical treatment; GS = Gleason score.

showed an association with the rate of change (pooled HR 1.39, 95% CI 1.15, 1.70; p = 0.001; Table 3).

3.5. Reasons for changing to radical treatment

Of the studies included, 22 reported the reasons for changing treatment [9,10,12,14–19,21–23,25–30,32–35]; note that the total for the following means adds to more than 100% because some studies only considered a subset of the possible reasons (Fig. 4). Up to an average of 38% of changes to radical treatments were because of a reclassified Gleason grade (95% CI 27%, 49%, Fig. 4) [14–19,21–23,25–29,33–35], while an average of 29% were caused by men meeting PSA, PSADT, or PSAv triggers in nine studies (95% CI 17%, 41%) [14–17,23,25,26,28,30,32,34,35]. Other tumor-related reclassifications such as the number of positive cores caused an average of 29% of the changes (95% CI 16%, 42%) [14,17,19,21–23,25,27,28,33–35]. In the 17 studies

![Fig. 4 – Forest plot of the proportion of patients undergoing radical treatment in studies who provided reasons for a change in treatment of (A) Gleason score increase, (B) PSA increase, (C) other biopsy reclassification, and (D) choice or anxiety. The year of initial recruitment is in parentheses after the study setting. Some confidence intervals have negative values because of small sample sizes and consequent invalid normal approximations. References for the study settings are provided in Table 1.](image-url)
determining reasons, an average of 20% of treatment changes were because of patient choice or anxiety (95% CI 14%, 27%) [12,14–19,22,23,25–27,29,30,32,34,35].

3.6. Quality of studies

Because the eligibility criteria varied so much among the studies, the cohorts of men with clinically localized prostate cancer also varied to greater and lesser degrees. There was no evidence to suggest that studies deviated from the given follow-up protocols or triggers for clinical review, although consistency would be increasingly difficult in the larger, multinational studies such as PRIAS [23]. Loss to follow-up was minimal, with almost all men accounted for as remaining on AS, changing to radical treatment, or dying during AS. Follow-up was <5 yr in 19 of the studies, with a median of 3.5 yr. This short-term follow-up of low-risk men is probably too short for an accurate estimate of longer-term clinical outcomes (such as metastases).

3.7. Bias in reporting

Plots of the percentage change to radical treatment against sample size and duration of follow-up are provided in Figure 5. There appears to be no convincing evidence of a relationship between cohort size and the estimated percentage of men changing to radical treatment, and hence no evidence of a small-study effect that could be due to publication bias. The large cohorts have a rate in the middle of the distribution of rates across the 26 cohorts. Figure 5 reveals a weak negative relationship between duration of follow-up and rate of radical treatment, suggesting that mature cohorts have a lower rate of changing to radical treatment.

3.8. Discussion

We reviewed 26 studies of men with clinically localized prostate cancer managed by AS and showed, as previously [6], that there is little consensus about inclusion criteria, follow-up protocols, or policy on triggers for change to radical treatment in contemporary programs. The meta-analysis revealed that there were eight prostate cancer deaths and five further cases of metastases in the overall cohort of 7627 men. Each year, 8.8% (95% CI 6.7%, 11.0%) of men on AS received radical treatment, with the rate of change varying from 11 to 222 per 1000 person-years, reflecting widely varying protocols. The rate of change to radical treatment was highest among men with higher baseline PSA levels and T-stage disease, and when programs scheduled repeat biopsies. Programs including men with Gleason >6 had fewer treatment changes. The most common reasons reported for changing to radical treatment were biopsy findings, PSA triggers, and patient choice. Although the reviewed studies had accrued 24 981 person-years of follow-up, the median follow-up was only 3.5 yr, and 19 studies had follow-up of <5 yr. This remains a serious limitation to the evidence base because most cases of progression in low-risk prostate cancer occur 10–15 yr after diagnosis.

The findings of this latest review and meta-analysis provide insights into the difficulties, uncertainties, and contradictions inherent in the contemporary management of localized prostate cancer using AS, although the limited and short-term evidence base makes it very difficult to reach clear conclusions about how AS programs should be designed in future. The primary reason for this is the lack of robust evidence: the two published treatment trials (SPCG-4 and PIVOT) [3–5] did not evaluate AS, but the more passive watchful waiting approach. Robust evidence will start to accrue with the publication of the outcomes of the ProtecT randomized controlled trial evaluating active monitoring in comparison with radical surgery and radiotherapy for men with clinically localized prostate cancer detected following PSA testing [37].

The concept of active surveillance or monitoring only emerged in the late 1990s, and considerable changes have occurred in the diagnosis and management of prostate cancer since then. In early cohorts, men presented clinically and were followed up with watchful waiting until PSA testing became more common [10,12]. During the 1990s and 2000s, the use of PSA testing facilitated more active programs through the use of PSA kinetics, but these varied widely as to whether they included men with low- or
intermediate-risk disease, and/or scheduled rebiopsy. The reported upward migration in the assignment of Gleason scores will also have affected the application of inclusion criteria [38,39]. More recently, increasing reluctance to rely on PSA kinetics has led to a greater interest in biopsy assessment developments, including the use of templates and magnetic resonance imaging (MRI) to provide enhanced risk profiles [40]. This has led to use of the full range of risk and biopsy criteria available in the most recently initiated studies to limit inclusion in AS programs to men with the lowest-risk disease, with protocols and triggers aimed at identifying progression at the earliest opportunity to facilitate change of management.

A key difficulty for AS programs has always been how to identify true disease progression. The rates of change from AS to radical treatment (8.8% per year in this meta-regression) is much higher than the levels of progression that should be expected from our knowledge of the natural history of prostate cancer [41]. In the SPCG-4 study [4,5], for example, with over 18 yr of follow-up of men with low-risk disease assigned to the more passive watchful waiting strategy, 14% died from prostate cancer and a large proportion did not require palliative care.

Several factors were associated with the rate of change to radical treatment in the meta-analysis. Studies with scheduled rebiopsy had the highest rate of treatment change, as expected, because rebiopsy offers the opportunity to detect small areas of high-grade tumor missed by initial biopsies, particularly among men with the most common baseline score of Gleason 3 + 3. Surgical series have shown that upgrading (as opposed to true progression) occurs at a rate of approximately 25–30% [42–44]. Higher baseline PSA and T stage were associated with a higher rate of change to radical treatment. PSA changes (including PSADT and PSAv) also led to many changes in treatment, although evidence of the validity of PSA as a measure of disease progression is lacking [45–48]. There was no evidence that any PSA eligibility criteria, frequency of PSA testing, or adoption of PSA triggers for treatment change were associated with rate of treatment change among studies, and only weak evidence that less mature studies had a higher rate of change.

Studies enrolling men with a Gleason score of 7 (3 + 4 or 4 + 3) had lower rates of treatment change compared to those with Gleason ≤6, perhaps because they used less stringent monitoring criteria. However, the pooled effect within the four studies including Gleason score 7 men showed a strong association between higher Gleason grade and an increased number of changes to radical treatment. Thus, although cohorts including men with Gleason 7 had a lower overall rate of radical treatment, the relatively small number of Gleason 7 men within these cohorts had a higher rate of switching to radical treatment. Many studies seek to enroll only men with the lowest risk disease and include aggressive monitoring strategies that require regular PSA testing and rebiopsies, leading to high rates of treatment change. Although this approach appears to have a high level of AS “failure”, this is advocated as the way to provide a very safe program, in which only a highly selected group of the lowest of low-risk patients is maintained. However, these cohorts represent only a small proportion of men with prostate cancer who will not suffer progression during their lifetimes [3–5].

Another reaction to the higher risk of men with Gleason 7 disease switching to radical treatment is to suggest improving monitoring strategies to better identify those men. However, although Gleason grade is a good prognostic factor, the association between a Gleason score 3 + 4 or 4 + 3 and subsequent clinical outcomes is uncertain, as very little is known about the natural history of these refined Gleason grades and the association with progression to clinically significant disease. Recent UK National Institute for Clinical Excellence guidelines have proposed a new consensus policy for AS in the UK to include an initial MRI with biopsy at around 1 yr so that men with larger tumors found on MRI or higher-grade tumors found on rebiopsy can be excluded from AS, or offered counseling of likely increased risks of radical treatment [40]. Research into imaging and the genetics of progression may offer further opportunities for identifying those most suitable for inclusion in or exclusion from AS. It has also been suggested that the uncertainties around the progression of Gleason 7 prostate cancer could be used as a justification for including intermediate- or higher-risk men in monitoring programs [49]. However, the current lack of understanding of the factors leading to true progression in men with Gleason 6 or 7 disease continues to hamper the development of consensus about how to define entry into and progression from AS.

On average, 20% of men changed from AS to radical treatment without meeting clinical triggers but because of patient anxiety or choice. The early stages of many AS programs clarify the disease risk characteristics by rebiopsy and frequent PSA measures, and upgrading of biopsies and the volatility of PSA kinetics can lead to perceptions of progression, resulting in anxiety and a wish to change to radical treatment. The relatively high number of men changing because of anxiety or choice reflects the uncertainties inherent in the monitoring of AS.

The strengths of this review include that the articles were found systematically and screening of potentially eligible papers was performed in duplicate. Any citation exclusions, data extractions and syntheses were checked by at least one other reviewer. Authors of publications were contacted with requests for further data in some cases and to check data extraction was correct. Conference abstract authors were contacted to provide full papers if they were available. However, although the methods of review were robust, the available data were very limited, particularly in terms of the length of follow-up. Further data, including definitive outcomes, are needed to clarify the longer-term outcomes and impact of AS. In the meta-regression, explicit information about men changing treatment was not available for all studies, so all those receiving radical treatment were grouped together. It would be beneficial in future publications if detailed reasons for treatment change were given, with clarity about the reason for leaving AS and whether men opt for radical treatment or
other options, including watchful waiting, particularly as they age.

4. Conclusions

The limited evidence on the long-term effectiveness of AS strategies and the wide variety of protocols followed mean that the clearest recommendations can be made in relation to research rather than clinical practice.

In terms of research, because such long follow-up is required, it would be helpful if the wide range of current AS cohorts continue to be followed up using their particular strategies, with improved collection of data on inclusion criteria, monitoring strategies, and reasons for changes from AS to other treatments. This will then permit future meta-analyses and modeling studies to reach conclusions about potentially effective and efficient strategies. It would also be helpful if these cohorts could be expanded (personal communication with R. Etzioni about the PROMIS study, 2014), or new cohorts initiated, as the number of men currently receiving AS remains small. It would be ideal if all men included in AS could also have contemporaneous research data collected, and if new randomized controlled trials of AS programs could be initiated, although these will take 10–15 yr to mature.

In the meantime, in terms of clinical practice, there remain uncertainties around AS, with no clear evidence as to which aspects of the many different eligible and monitoring protocols are most important for long-term outcomes. In this review, we have identified some individual characteristics related to the likelihood of patients continuing on AS, including baseline PSA, T stage, and Gleason grade. This information will enable those overseeing AS programs to identify the men most likely to change treatment. We also identified protocol characteristics (eg, scheduled biopsies) related to the overall proportion of patients changing treatment. This information will help in designing AS protocols to minimize the rate of treatment change. The high rate of non-protocol-related change found here suggests that AS programs should increase the provision of reassurance and counseling for AS patients. AS protocols should be designed and updated through an evidence-based process that takes into account evolving diagnostic and prognostic technologies, as well as our understanding of the biology underlying prostate cancer, to maximize safety and quality of life for patients while avoiding overtreatment of disease.

Ultimately, what is needed is robust and long-term evidence to convince clinicians and patients of the safety of AS [37]. Meanwhile, clinicians and patients will have to continue to make decisions about AS based on their own interpretations of the limited evidence base and perceptions of the balance of risks and benefits in the management of localized prostate cancer.

Study concept and design: Simpkin, Metcalfe, Martin, Tilling.
Acquisition of data: Simpkin.
Analysis and interpretation of data: Simpkin, Metcalfe, Martin, Tilling.
Drafting of the manuscript: Simpkin, Donovan.
Critical revision of the manuscript for important intellectual content: Metcalfe, Martin, Lane, Donovan, Hamdy, Holmberg, Neal, Tilling.
Statistical analysis: Simpkin.
Obtaining funding: None.
Administrative, technical, or material support: None.
Supervision: Tilling, Metcalfe, Martin.
Other (specify): None.

Financial disclosures: Andrew Simpkin 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: Andrew Simpkin is funded by the Medical Research Council (MR/L011824/1) and previously by the NIHR Health Services and Delivery Research program (project number 09/2000/63), Jenny L. Donovan, Freddie C. Hamdy, and David E. Neal are the Principal Investigators and J. Athene Lane is the Trial Coordinator of the Prostate Testing for Cancer and Treatment (ProtecT) trial. Richard M. Martin, Jenny L. Donovan, Freddie C. Hamdy, and David E. Neal are the Principal Investigators of the Comparison Arm to ProtecT (CAP) trial. The ProtecT trial is funded by the UK National Institute for Health Research, Health Technology Assessment Programme (HTA 96/20/99) and the CAP trial by Cancer Research UK/UK Department of Health (C11043/A4286, C18281/A8145 and C18281/A11326). Funding for additional research has been received from the World Cancer Research Fund, the University of Bristol Cancer Research Fund and the National Cancer Research Institute (formed by the Department of Health, Medical Research Council and Cancer Research UK).

Funding/Support and role of the sponsor: This project was funded by the NIHR Health Services and Delivery Research Programme (project number 09/2000/63) and will be published in full in the NIHR Journals Library. Further information available at http://www.journalslibrary.nihr.ac.uk/hsdr

This report presents independent research commissioned by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the Health Services and Delivery Research programme, NIHR, NHS or the Department of Health.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2015.01.004.

References

Clinical outcome of patients with...N Engl J Med 2014;370:932–42.


