Title: Defining and measuring adherence in observational studies assessing outcomes of real-world active surveillance for prostate cancer: a systematic review

Authors:
Glenda Kith, MSBH\textsuperscript{ab}, Sarah Lisker, BA\textsuperscript{bc}, Urmimala Sarkar MD, MPH\textsuperscript{bc}, Jill Barr-Walker MS, MPH\textsuperscript{d}, Benjamin N. Breyer, MD, MAS, FACS\textsuperscript{ef}, Nynikka R. Palmer DrPH, MPH\textsuperscript{bce}

† Glenda Kith and Sarah Lisker should be considered joint first author.

Corresponding Author:
Nynikka R. Palmer DrPH, MPH
Assistant Professor
Division of General Internal Medicine at Zuckerberg San Francisco General Hospital
University of California San Francisco
1001 Potrero Avenue, Box 1364
San Francisco, CA 94143
Telephone: 415-206-4334
Fax: 415-206-5586

\textsuperscript{a} San Francisco Department of Public Health
\textsuperscript{b} Department of Medicine, Zuckerberg San Francisco General Hospital and Trauma Center, University of California San Francisco
\textsuperscript{c} Center for Vulnerable Populations, University of California San Francisco
\textsuperscript{d} ZSFG Library, Zuckerberg San Francisco General Hospital and Trauma Center, University of California San Francisco
\textsuperscript{e} Department of Urology, University of California San Francisco
\textsuperscript{f} Department of Epidemiology and Biostatistics, University of California San Francisco
E-mail: Nynikka.Palmer@ucsf.edu

**Keywords:** clinical protocols; guideline adherence; monitoring, ambulatory; patient safety; prostatic neoplasms

**Word Count (Text):** 3,171

**Word Count (Abstract):** 287
Abstract

Context: Evidence-based guidelines for active surveillance (AS), a treatment option for men with low-risk prostate cancer, recommend regular follow-up at periodic intervals to monitor disease progression. However, gaps in monitoring can lead to delayed detection of cancer progression, leading to missed window of curability.

Objective: We aimed to identify the extent to which real-world observational studies reported adherence to monitoring protocols among prostate cancer patients on AS. When reported, we sought to characterize definitions of adherence.

Evidence Acquisition: We systematically reviewed observational studies assessing outcomes of prostate cancer patients on AS, published before March 22, 2019 in PubMed, Embase, and CENTRAL. Adherence definitions were considered time-bound if they included a pre-specified time, and binary if adherence was assessed but did not specify a time interval. We assessed study quality using the Strengthening the Reporting of Observational Studies in Epidemiology checklist.

Evidence Synthesis: Forty-five studies met our inclusion criteria. Eleven studies did not report any data on adherence to AS protocols. Twenty-five studies did not explicitly measure adherence, but provided relevant data (e.g. number of patients who received a repeat biopsy). Six studies reported adherence using a time-bound definition, while three studies used a binary definition. Twenty-three studies provided information on patients lost to follow-up.
Conclusion: Most studies reporting outcomes of patients on AS did not measure or report adherence. When reported, adherence was often not time-specific. As some AS patients will benefit from maintaining a window of curability, clinical practices and future studies should track and report adherence and associated factors.

Patient Summary: We reviewed real-world observational studies examining outcomes of prostate cancer patients on AS. Most studies did not clearly define or report adherence to monitoring protocols, which is important to consider for appropriate disease management.
Introduction

Prostate cancer is the most common cancer in men in the United States, and second most common internationally.\(^1,2\) Most prostate cancers are low-risk and slow growing; however, a subset can progress to more lethal disease.\(^3\) Active surveillance (AS) is the recommended standard of care for low- and some intermediate-risk prostate cancers that have little risk of progression.\(^4\) AS defers active treatment (e.g., surgery or radiation) with close monitoring and leads to active treatment if the cancer progresses (e.g., low grade disease becomes higher grade disease), the cancer is reclassified (e.g., detected a previously undetected cancer of higher grade on confirmatory or surveillance biopsy), or the patient decides to pursue active treatment. AS also provides the opportunity to maintain men’s quality of life without the potential adverse effects of treatment.\(^5\)

AS eligibility criteria and monitoring guidelines vary widely by institution, but generally recommend regular monitoring for disease progression that includes prostate-specific antigen (PSA) tests, digital rectal exams (DRE), and repeat prostate biopsies at periodic intervals.\(^6\) However, the extent of patient adherence to AS protocols is unclear. Little is known about which populations or settings are more vulnerable to monitoring gaps, or multi-level factors associated with non-adherence or loss to follow-up (LTFU).

In this systematic review, we sought to identify if and how real-world observational studies tracked and reported adherence to AS protocols when evaluating outcomes of prostate cancer patients on AS. Specifically, we aimed to assess whether or not such studies reported patient adherence to AS protocols (e.g., PSA tests, DREs, repeat biopsies, and duration and
frequency of monitoring received). When studies reported adherence, we sought to understand how investigators define and measure adherence, as well as the range of adherence rates.

Evidence Acquisition

Our review was guided by the PRISMA statement. We conducted a systematic literature search (PROSPERO #CRD42016051128) on March 8, 2018 and updated on March 22, 2019 in PubMed, Embase, and CENTRAL. We created our search strategy in collaboration with a clinical librarian (JBW) using text words and controlled vocabulary, including MeSH and Emtree terms (e.g., “active surveillance” and “Prostatic Neoplasms”[Mesh]). Appendix A details the complete search strategy.

Study eligibility criteria included: 1) focus on AS for prostate cancer; 2) real-world observational study reporting actual patient care (e.g., a retrospective medical chart review); 3) followed a predetermined evidence-based AS clinical eligibility criteria (e.g., PSA level, Gleason score, stage, number of positive cores, etc.) and an AS monitoring protocol (e.g., National Comprehensive Cancer Network [NCCN] guidelines – PSA every 6 months, annual DRE and biopsy); and 4) reported outcome data (e.g., survival rates, movement to active treatment, or LTFU). We excluded non-English studies, studies that evaluated outdated AS protocols, as well as studies that solely reported on active treatment patients. Clinical trials, interventional studies, efficacy trials (e.g., study cohorts with a study protocol), or studies comparing different treatment modalities to test protocols or treatment effectiveness were also excluded.

Two reviewers (GK, SL) independently screened titles and abstracts to determine relevancy and eligibility, and completed full text reviews. A third reviewer (US, BB, or NP)
reconciled discrepancies. We collected several data elements on adherence and outcomes (Table 1). We categorized studies as reporting adherence if they explicitly included a measure with the term “adherence” or “compliance”. We considered the adherence/compliance definition time-bound if it included the timeliness of monitoring (e.g. patient adherent if repeat biopsy completed within six months of a designated date). We considered the definition binary if it assessed adherence, but was not time-bound (e.g. testing received versus not). We tabulated adherence definitions and results, when provided. We used the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist to assess study quality.8

Evidence Synthesis

We screened 6,118 articles’ titles and abstracts and reviewed the full text of 267 articles. After applying our inclusion and exclusion criteria, we included 45 studies (Figure 1), and classified them into four mutually exclusive groups (Figure 2).

Study characteristics

After removing participant sub-groups who did not meet our eligibility criteria (e.g., enrolled in a randomized controlled trial or who did not undergo AS), we ultimately included a total of 29,143 participants (range: 34 – 5,302) from 45 studies. The earliest enrollment in an AS program began in 1990,9,10 culminating as late as 2017.11

Eligibility criteria into AS programs varied across studies; however, most studies restricted inclusion by diagnostic criteria – e.g., Gleason score ≤ 6, PSA ≤ 10 ng/ml, ≤ 33% positive biopsy cores, and ≤ 50% involvement in a single core. Some studies stated they included select men with higher risk features, such as those with a strong preference for AS.9,12–
A few studies considered patients’ cognitive capacity to understand AS programs as well as their ability to attend follow-up visits. Eligibility criteria also changed over time in certain long-term studies, reflecting a general trend in the evolution of AS protocols. Most studies reported follow-up time on AS, however the measure varied based on when investigators determined follow-up began (e.g. upon diagnosis or agreement to enroll into an AS program) and ended (e.g. upon study completion, movement to active treatment, or other censorship). Median follow-up time ranged from 16.9 months to 6.7 years.

Patient characteristics

All 45 studies reported patient age and clinical characteristics – such as Gleason score, PSA, number of positive biopsy cores, risk classification, and clinical stage. Fourteen studies reported patient race/ethnicity. Four of these publications specifically focused on studying racial/ethnic minorities. Two studies reported patients’ marital status. Two studies also reported tobacco use history; however, only one also reported history of substance abuse, mental illness, homelessness, and primary language.

AS protocols

Follow-up protocols varied across studies, using strict or modified NCCN, Prostate Cancer Research International Active Surveillance (PRIAS), Johns Hopkins, University of California San Francisco, Epstein, and other institution-specific protocols. All studies described PSA testing and biopsy schedules. Nine studies followed a protocol to assess PSA every three months for the first two years on AS, followed by every six months thereafter. For all other studies, PSA testing
intervals ranged from one to 12 months. Repeat biopsy schedules varied widely across studies as seen in seven publications that biopsied patients annually,\textsuperscript{17,18,35,40–42,45} and five that biopsied patients at one, four, seven, and ten years, and every five years thereafter.\textsuperscript{25,26,43,46,47} Follow-up protocols also varied by risk level at diagnosis and during follow-up, patient preference, and patient age. As with eligibility criteria, protocols changed over time to accommodate evolving evidence about AS safety and effectiveness.

**Reported Adherence**

While 11 studies did not report any data on adherence to AS protocols,\textsuperscript{11–13,18–20,31,32,35,37,45} 34 studies reported some form of adherence (e.g. time-bound, binary, or undefined).

**Time-bound Definitions of Adherence**

Six studies reported adherence using a time-bound definition.\textsuperscript{23,25,40,46,47,49} We considered adherence definitions time-bound if investigators' definitions included the number of monitoring events within a specific time interval. Reported adherence rates are included in Table 2.

No studies had the same adherence definition, although some were similar with varying degrees of flexibility. For example, Osterberg et al considered patients adherent if they received a biopsy within 12 months of the scheduled date,\textsuperscript{23} whereas Tosoian et al, Bul et al, Bokhorst et al, and Soeterik et al determined patients to be adherent if they received biopsies within six months of the scheduled date.\textsuperscript{25,40,46,47} Lee et al did not provide additional time beyond the recommended interval for a repeat biopsy.\textsuperscript{49}

**Binary Definitions of Adherence**
Three studies used binary definitions of adherence that categorized surveillance as complete or incomplete, but did not include a specific time interval in which monitoring events needed to take place.\textsuperscript{30,33,41} For example, Becker et al assessed the proportion of patients that received four scheduled repeat biopsies.\textsuperscript{41} Similarly, Sugimoto et al evaluated the proportion of patients that received their first scheduled repeat biopsy.\textsuperscript{33} And Hefermehl et al examined the proportion of PSA tests and biopsies performed compared to the total that should have been completed.\textsuperscript{30} We summarize these results in Table 3.

**Undefined measures of adherence**

We included 25 studies that did not explicitly define adherence, but either (1) reported descriptive statistics of follow-up testing rates, including a time frame (such as median time from diagnosis to first repeat biopsy), or (2) included an incomplete or nonspecific measure of adherence (such as the absolute number of men who received repeat biopsies, rather than a proportion).\textsuperscript{9,10,14–17,21,22,24,26–29,34,36,38,39,42–44,48,50–53}

Twelve studies described the number of follow-up events within a specified time interval.\textsuperscript{15,17,21,22,27–29,34,42,44,51,52} Common measures included the time between surveillance exams and proportion of men who received a follow-up test within a time frame. For example, Hashine et al reported 55.2\% of their cohort received the first repeat biopsy after a median 1.1 years on AS.\textsuperscript{29} Meunier et al reported a mean of 1.0 ± 0.49 years between repeat biopsies.\textsuperscript{22}

The remaining 13 studies did not define adherence but reported a measure that we considered representative of adherence.\textsuperscript{9,10,14,16,24,26,36,38,39,43,48,50,53} For example, some studies reported the median number of biopsies received\textsuperscript{10,14,38} and the number of men who received a follow-up test.\textsuperscript{9,14,16,24,26,36,38,48,50,53} However, these measures lacked a corresponding time
frame, such that it is unclear if the tracked biopsies occurred within one year, two, or throughout the entire study period. Other studies lacked key information, such as the total number of patients on AS to compare to the number that received follow-up testing.26,50,53

**Lost to Follow-Up (LTFU)**

Twenty-three studies provided information on patients LTFU. Rates of LTFU vary from none52 to 19.5%30, however this measure was not consistently defined across studies making comparison infeasible. We saw various approaches to defining whether or not a patient was LTFU. Hefermehl et al considered patients LTFU after no response or information about the patient more than 12 months after the last scheduled appointment.30 Osterberg et al, alternatively, considered 31% of patients LTFU after failing to reach them after three attempted phone calls, until querying a cross-institutional database that indicated only 17.3% of patients were truly LTFU.23 In addition, some studies excluded patients LTFU.10,39,40,42

Ballas et al examined the association between the chance of being LTFU and socioeconomic status (SES).35 They found that the cumulative incidence of being LTFU at a safety-net hospital was significantly higher (57% at 5 years) than at a comprehensive cancer center (37% at 5 years), and that lower SES was associated with likelihood of being LTFU.

**Study Quality and Potential Bias**

According to the STROBE checklist,8 the overall quality of studies included was high. Across most studies, we found limited reporting on confounders related to patient outcomes, including clearly defined measures of adherence and patient characteristics such as SES, race, ethnicity, primary language, marital status, and other characteristics shown to influence access to care. Additionally, not all studies provided details on efforts to address potential sources of
or addressed how investigators handled missing data.\textsuperscript{9,10,12,14,15,17,18,20–22,24,25,27,28,31,33,34,37,40,43,44,46,47,49–53}

Discussion

The majority of observational studies evaluating clinical outcomes of men on AS do not measure adherence. When adherence is reported, assessment across studies and settings is heterogeneous. Not all studies provided a definition of the measure that considers the timeliness of the routine testing. In addition, a few studies excluded patients LTFU. Since the premise of AS is routine scheduled monitoring to assess disease progression or reclassification, timely adherence is essential to retain a window of curability. Although some studies report no significant mortality difference between AS and active treatment among men with low-risk prostate cancer,\textsuperscript{54,55} adherence to monitoring remains paramount for long-term patient safety, particularly among a subgroup of patients. Up to 30\% of men diagnosed with low- and intermediate-risk prostate cancer harbor more aggressive disease missed on biopsy,\textsuperscript{47,56–58} and some men with higher risk disease may still opt for AS that necessitates regular follow-up.\textsuperscript{4} African American men may have a higher risk of disease reclassification, as found in a recent systematic review.\textsuperscript{59} Additionally, men with intermediate-risk disease who can be safely monitored on AS have more than a three times higher risk of metastasis at 15-year follow-up.\textsuperscript{57,60} AS – including the extent to which it is followed - should be reported and studied to further understand the quality of AS care, resources needed to ensure high-quality monitoring (e.g., registries), and how protocols can be more widely implemented across diverse settings and populations safely to prevent overtreatment without compromising quality of life and
survival. This involves understanding a range of constructs including adherence and various implementation outcomes, which is critical to understanding the wider implementation and dissemination of AS in clinical practice.

Most studies did not evaluate the potential impact of multi-level factors on adherence. For example, only one study examined whether or not patients on AS experienced homelessness or had a history of mental illness, and only two studies detailed how clinics performed monitoring or responded to patients’ concerns surrounding AS (e.g. a dedicated nurse conducted follow-up activities and supported patients experiencing anxiety). Prior research indicates patient-, provider-, and system-level factors such as these may influence adherence. For example, Kinsella et al identified multi-level barriers and facilitators to AS adherence including cancer characteristics (e.g., lower PSA), patients’ perceived benefits and harms, social support (e.g., peer support, partner anxiety), receipt of useful information communicated by providers, healthcare organization and practice (e.g., educational classes on AS), and health policy factors (e.g., national guidelines). We also recognize that adherence to AS protocols may vary by patient preference (e.g., refuses repeat biopsy), physician practice (e.g., tailors based on clinical and follow-up characteristics), health care system and country (e.g., follow NCCN versus the American Urological Association [AUA] guidelines or others), and the availability of multi-parametric magnetic resonance imaging. Identifying barriers and facilitators to AS adherence is crucial for successful long-term disease management that delays or avoids unnecessary treatment until it is warranted, and can inform intervention development.
As a framework to examine the interrelated components affecting high quality care, including adherence, the Chronic Care Model is a promising model for improving management of chronic conditions. AS constitutes ongoing care for a chronic condition; as such, this framework suggests looking to a health system situated within a community, in which productive interactions between a prepared, proactive practice team and engaged and active patients drive positive health outcomes. Patient-centered strategies such as primary care provider (PCP) and specialty provider coordination, between-visit self-management support, technology-enabled workflows, as well as organizational support of providers monitoring high-risk populations are promising avenues for safe and effective long-term care. Such team-based care and technology interventions present specific opportunities to improve the safety of outpatient monitoring. For example, PCPs can help ensure appropriate monitoring for their patients by ordering follow-up tests in collaboration with urology, and electronic medical record-based alerts and registries can prompt monitoring activities when well-integrated into care team workflows. If AS protocols are meant to be effective for detecting prostate cancer progression, guidelines must incorporate clear and defined adherence measures while offering care providers and patients guidance on tracking and following monitoring schedules. As experts call for consensus on AS eligibility criteria, monitoring guidelines, and thresholds for intervention, adherence must be measured and considered in the evaluation and selection of an optimal AS protocol.

Future studies should track and report variables relevant to follow-up care for and adherence to AS protocols. First, studies should determine a definition of adherence to AS protocols based on guideline-recommended time (e.g., every 6 months) or a pre-defined grace
period (e.g. within three months of recommended testing). Second, an adherence rate to
monitoring (e.g., completed PSA testing, annual biopsy and DRE, etc.) should be calculated,
along with the average time between monitoring. Third, studies should also report AS
outcomes, including number and rate patients moved to active treatment, whether due to
disease progression, reclassification, or patient preference. Lastly, studies should define and
assess LTFU, including associated factors (e.g., race, comorbidities), to identify patients at
increased risk.

Even with clear recommendations for monitoring, improving adherence in real-world
clinical settings can be challenging. Fragmented health records and limited resources in safety-
net settings can introduce unique barriers to implementing evidence-based recommendations
in prostate cancer care, particularly for AS. Patient-centered and technology-enabled
solutions to measure and improve adherence must also be designed for and tested in settings
that disproportionately care for low-income and vulnerable populations, many of whom
already face barriers to accessing care.

While we report on studies that include both measures of adherence and compliance, it
is important to differentiate between the two terms, as recommended by the World Health
Organization. The former connotes an active agreement between patient and provider: does
the patient adhere to an agreed-upon recommendation? The latter suggests a more passive
role of the patient: does the patient comply with a physician’s orders? When discussing AS, the
success of which is dependent on the patient’s agreement and understanding of follow-up care,
we believe that “adherence” better reflects the inherent shared decision-making paradigm.

There are limitations to this systematic review. First, since AS protocols differ across
settings, we could not compare adherence across studies. As a result, we decided to focus on primary data analysis of real-world observational studies within single or multiple settings using the same follow-up protocol, rather than secondary data analysis of cancer registries or consortiums employing various protocols. We also limited our review to studies that had predetermined evidence-based AS clinical eligibility criteria and monitoring protocols, which excluded a few studies (e.g., those reporting on the MUSIC consortium).\textsuperscript{73,78} This was done to capture studies that intended to follow patients who were clinically appropriate and monitored based on a predetermined protocol (e.g., NCCN). We also included non-US studies and recognize that differences in health care systems across countries have implications for variations in access to and quality of care. However, we hope our review illustrates the need for consistent definitions and reporting of adherence in order to move towards a standardized AS protocol. Also, the causality between adherence to AS and patient outcomes is not well understood due to the focus on short-term outcomes in the literature; however, successful receipt of timely follow-up visits and testing should be treated as a necessary metric when reporting AS outcomes.\textsuperscript{61} To our knowledge, this is the first systematic review that assesses the extent adherence to follow-up protocols is measured in real-world observational studies reporting outcomes of men on AS. Due to the large number of men managed by AS, this work offers valuable implications that may improve the delivery of care for prostate cancer patients on AS. In addition, previous studies demonstrating differences in prognostic value and prostate cancer specific mortality for men on AS by race and ethnicity indicate that it is imperative to thoroughly understand factors associated with successful disease management, including
adherence, in order to move the needle towards recognizing and reducing disparities.\textsuperscript{79,80}

**Conclusions**

Adherence to evidence-based AS protocols is not uniformly reported or defined in the scientific literature. As AS becomes a more common management strategy among men with low-risk prostate cancer, our findings highlight the need for more research to: (1) establish adherence definitions and measures relevant to AS; (2) include adherence as a quality improvement measure for AS outcomes; and (3) track and identify multilevel factors associated with adherence and LTFU.
Funding: Glenda Kith was supported by the National Cancer Institute of the National Institutes of Health under award number 5R25CA078583. Dr. Sarkar is supported through grants from the Agency for Healthcare Research and Quality (P30HS023558) and the National Cancer Institute of the National Institutes of Health under award number K24CA212294. Dr. Palmer is supported by the National Cancer Institute of the National Institutes of Health under award number K01CA211965. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality or the National Institutes of Health. The funders had no role in the design or conduct of the study, collection, management, analysis, or interpretation of data, nor the preparation, review, or approval of the manuscript.

Conflict of Interest Disclosures: None.

Authors Contributions: GK and SL designed the study with support from US, BB, and NP, developed the search strategy with support from JBW, screened and extracted all records, and drafted the manuscript. US, BB, and NP resolved discrepancies during the review. JBW performed the database search, as well as advised on data collection and management practices. US conceived of the study topic in close partnership with NP. NP takes responsibility for the integrity of this work. GK and SL served as co-first authors, each with equal contribution to the manuscript. All authors critically reviewed the manuscript and approved the final version, and meet ICMJE criteria for authorship.
Acknowledgements: We would like to thank Gem Le for her mentorship, as well as Karen Kwaning and Deepak Maharaj for their support with data collection.

Data Access and Responsibility: Dr. Palmer 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.

Data Sharing: Data are available upon request.

PROSPERO Registration: PROSPERO 2016:CRD42016051128

Registration Link:

http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016051128
RE utens:


Figure legends:

Figure 1: PRISMA Flow Diagram

Figure 2: Flow diagram categorizing included studies
### Table 1: Data Elements

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Design</td>
<td>Setting (e.g. private hospital, safety-net, population-based)</td>
</tr>
<tr>
<td></td>
<td>Number of participants</td>
</tr>
<tr>
<td></td>
<td>Country</td>
</tr>
<tr>
<td></td>
<td>Active surveillance (AS) eligibility requirements (e.g. age, tumor grade, prostate-specific antigen [PSA] level)</td>
</tr>
<tr>
<td></td>
<td>AS monitoring protocol (including testing type and frequency of monitoring)</td>
</tr>
<tr>
<td>Study Group</td>
<td>Age</td>
</tr>
<tr>
<td>Demographics</td>
<td>Race/Ethnicity</td>
</tr>
<tr>
<td></td>
<td>Comorbidities</td>
</tr>
<tr>
<td></td>
<td>Social demographics (e.g. Social economic status, homelessness, insurance status)</td>
</tr>
<tr>
<td>Adherence</td>
<td>Definition of adherence as defined by the study</td>
</tr>
<tr>
<td></td>
<td>Any measure of patients adhering to PSA testing schedule</td>
</tr>
<tr>
<td></td>
<td>Any measure of patients adhering to biopsy schedule</td>
</tr>
<tr>
<td></td>
<td>Any measure of patients adhering to digital rectal exam schedule</td>
</tr>
<tr>
<td>Outcome</td>
<td>Definition of lost to follow-up (LTFU) as defined by the study</td>
</tr>
<tr>
<td></td>
<td>Any measure of patients LTFU</td>
</tr>
<tr>
<td></td>
<td>Any measure of patients moved to watchful waiting</td>
</tr>
<tr>
<td></td>
<td>Any measure of patients moved to active treatment</td>
</tr>
<tr>
<td></td>
<td>Any measure of prostate cancer-specific death</td>
</tr>
</tbody>
</table>

AS = active surveillance; PSA = prostate-specific antigen; LTFU = lost to follow-up
Table 2: Characteristics of studies measuring adherence using a time-bound definition (n=6)

<table>
<thead>
<tr>
<th>Study info</th>
<th>Cohort name</th>
<th>Study period</th>
<th>Number of patients who elected Active Surveillance (AS)</th>
<th>Adherence measure</th>
<th>Adherence definition</th>
<th>Adherence results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tosoian et al. 2011</td>
<td>Johns Hopkins</td>
<td>1995-2010</td>
<td>796</td>
<td>Adherence to repeat biopsies</td>
<td>Patient undergoes biopsy within 0.5 years of designated date</td>
<td>Year 1: 92% adherent Year 2: 91% adherent Total (12 years): 89% (79-100%)</td>
</tr>
<tr>
<td>Bul et al. 2013</td>
<td>PRIAS</td>
<td>2006-2012</td>
<td>2494</td>
<td>Adherence to first repeat biopsy</td>
<td>Patient undergoes biopsy within 1.5 years of initial diagnosis</td>
<td>81% adherent</td>
</tr>
<tr>
<td>Bokhorst et al. 2015</td>
<td>PRIAS</td>
<td>2006-2014</td>
<td>4547</td>
<td>Adherence to scheduled prostate-specific antigen (PSA) tests</td>
<td>At least 3 PSA tests/year in first 2 years, at least 1 PSA test/year in following years.</td>
<td>Total: 91% adherent</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adherence to repeat biopsies</td>
<td>Biopsy within 6 months of designated date, or within 12 months for patients with prostate-specific antigen doubling time (PSA-DT) (3-10 years)</td>
<td>Year 1: 81% Year 4: 60% Year 7: 53% Year 10: 33% Total: 70% For men with PSA-DT(3-10): Year 2: 24% Year 8: 9% For men with &gt; 4 years of follow-up, compliance is 30% Year 1: 100%</td>
</tr>
<tr>
<td>Lee et al. 2010</td>
<td>Kansas City Veterans Affairs</td>
<td>2004-2009</td>
<td>45</td>
<td>Adherence to first PSA</td>
<td>Patient receives 1 of 4 recommended PSA within first year</td>
<td>Year 1: 53.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adherence to first repeat biopsy</td>
<td>Patient receives 1 biopsy within 1 year</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Proportion of patients that received follow-up PSA</td>
<td>Patient undergoes PSA every 3-6 months</td>
<td>≥1 follow-up PSA: 92% Median number of PSA tests: 7 (range 1-21). Year 10: &lt;10% Average number of biopsies: 2 (range 1-5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Proportion of men that received biopsy every 12-24 months</td>
<td>Patient undergoes biopsy every 12-24 months</td>
<td></td>
</tr>
<tr>
<td>Osterberg et al. 2017</td>
<td>UCSF</td>
<td>2004-2014</td>
<td>104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Abbreviations</td>
<td>Year 10: &lt;10%</td>
<td>Patient receives ≥ 75% recommended PSA tests for their follow-up duration</td>
<td>Total: 74% adherent (ranged from 55%-83% by hospital)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>---------------</td>
<td>--------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soeterik et al. 2019</td>
<td>PRIAS 2008-2014 958</td>
<td>Proportion of patients that received recommended PSA tests</td>
<td>Patient undergoes biopsy within 6 months</td>
<td>Year 1: Ranged from 48%-79% by hospital</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proportion of patients that underwent repeat biopsy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proportion of patients who received all scheduled biopsies</td>
<td>Patient who should undergo ≥ 1 biopsy misses none within follow-up period</td>
<td>Total: 52% adherent (ranged from 40%-68% by hospital)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** AS = active surveillance; PRIAS = Prostate Cancer Research International: Active Surveillance study; PSA = prostate-specific antigen; PSA-DT = prostate-specific antigen doubling time
Table 3: Characteristics of studies measuring adherence using a binary definition (n=3)

<table>
<thead>
<tr>
<th>Study info</th>
<th>Cohort name</th>
<th>Study period</th>
<th>Number of patients who elected Active Surveillance (AS)</th>
<th>Adherence measure</th>
<th>Adherence results</th>
</tr>
</thead>
</table>
| Becker et al. 2014\(^1\) | HAROW | 2008-2012 | 387/748\(^1\) | Proportion that receive repeat biopsies | Biopsy 1: 73.6%  
Biopsy 2: 45.2%  
Biopsy 3: 23.8%  
Biopsy 4: 11.4% |
| Sugimoto et al. 2015\(^3\) | PRIAS-JAPAN | 2010-2013 | 386 | Proportion that receive first repeat biopsy | Biopsy 1: 75.8% |
| Hefermehl et al. 2016\(^4\) | Kantonsspital Baden, Switzerland | 1999-2013 | 157 | Proportion of performed prostate-specific antigen (PSA) tests to total expected  
Proportion of performed biopsies to total expected | 39.7% of PSA values missing  
81% men received confirmation biopsy, 19% did not  
36.5% biopsies were missing |

Abbreviations: AS = active surveillance; HAROW = Hormonal treatment, Active Surveillance, Radiation therapy, OP, Watchful waiting study; PRIAS = Prostate Cancer Research International: Active Surveillance study; PSA = prostate-specific antigen

\(^1\)Subset of AS patients monitored by an AS protocol that met our inclusion criteria
Figure 1: PRISMA Flow diagram

- Records identified through database searching (n = 8023)
- Records identified from other sources (n = 2)
- Records after duplicates removed (n = 6118)
- Records excluded (n = 5851)
- Records screened (title/abstract) (n = 6118)
- Full-text articles assessed for eligibility (n = 267)
- Studies included in synthesis (n = 45)
- Full-text articles excluded, with reasons (n = 222)

- 97 Interventional, not representative of real-world practice
- 57 Used follow-up protocol not currently accepted
- 21 No eligibility requirements for entry into surveillance
- 14 No outcomes of interest included
- 10 Not published in English
- 9 Study does not report on active surveillance
- 8 No low-risk patients included
- 4 No predetermined follow-up protocol
- 1 Conference paper
- 1 Reprint of an earlier study
Figure 2: Flow diagram categorizing included studies

Included studies
(n = 45)

- Study does not measure adherence to active surveillance monitoring protocol (n = 11)
- Terms “adherence” or “compliance” are not used, but study provides a measure that is representative of adherence (n = 25)
- Terms “adherence” or “compliance” are defined (n = 9)
  - Binary definition (tracks monitoring activities, but no time interval specified) (n = 3)
  - Time-bound definition (tracks monitoring activities within a specified time interval) (n = 6)
Appendix A

No date limits. No language limits.

PubMed

("Prostatic Neoplasms"[Mesh] OR “Prostate Neoplasms” OR “Prostate Neoplasm” OR prostate neoplasm OR prostate neoplasms OR “Prostatic Neoplasm” OR “Prostate Cancer” OR “Prostate Cancers” OR “Cancer of the Prostate” OR “Prostatic Cancer” OR “Prostatic Cancers” OR “Cancer of Prostate” OR pca)
AND ("Watchful Waiting"[Mesh] OR “watchful waiting” OR “active monitoring” OR “watchful observation” OR “Active Surveillance” OR “surveillance” OR “Expectant Management” OR “Conservative Management”)
AND ("low risk” OR low risk OR “early stage” OR “favorable risk” OR favorable risk OR favourable risk OR indolent OR “clinically localized” OR localized OR localised OR “clinically insignificant” OR “low grade” OR “lower grade”)

Initial search on 11/3/16= 2082
Updated search on 3/8/18= 406
Updated search on 3/22/19 = 275

Embase

('prostate cancer'/exp OR 'prostate neoplasms' OR 'prostate neoplasm' OR 'prostatic neoplasm' OR 'prostate cancer' OR 'prostate cancers' OR 'cancer of the prostate' OR 'prostatic cancer' OR 'prostatic cancers' OR 'cancer of prostate' OR pca OR (prostate AND neoplasms)) AND ("watchful waiting" OR 'active monitoring' OR ‘watchful observation’ OR 'active surveillance' OR 'surveillance' OR 'expectant management' OR 'conservative management') AND ("low risk" OR 'early stage' OR 'favorable risk' OR indolent OR 'clinically localized' OR localized OR localised OR 'clinically insignificant' OR 'low grade' OR 'favourable risk' OR 'lower grade' OR (favorable AND risk) OR (favourable AND risk) OR (low AND risk))

Initial search on 11/3/16= 3550
Updated search on 3/8/18= 739
Updated search on 3/22/19 = 686

CENTRAL

("Prostatic Neoplasms" OR “Prostate Neoplasms” OR “Prostate Neoplasm” OR prostate neoplasm OR prostate neoplasms OR “Prostatic Neoplasm” OR “Prostate Cancer” OR “Prostate
Cancers” OR “Cancer of the Prostate” OR “Prostatic Cancer” OR “Prostatic Cancers” OR “Cancer of Prostate” OR pca)

AND (“watchful waiting” OR “active monitoring” OR “watchful observation” OR “Active Surveillance” OR “surveillance” OR “Expectant Management” OR “Conservative Management”)

AND (“low risk” OR low risk OR “early stage” OR “favorable risk” OR favorable risk OR favourable risk OR “favourable risk” OR indolent OR “clinically localized” OR localized OR localised OR “clinically insignificant” OR “low grade” OR “lower grade”)

Initial search on 11/3/16= 134
Updated search on 3/8/18= 94
Updated search on 3/22/19 = 57

All

Total = 8023
Duplicates= 1907
Final total = 6116