Ontario
Cancer Care Ontario
Action Cancer Ontario

Guideline 27-3 IN REVIEW

A Quality Initiative of the Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)

Magnetic Resonance Imaging for Pre-Treatment Local Staging of Prostate Cancer

M Haider, J Salerno, A Finelli, C Morash, S Morgan, N Power, N Schieda, and the Magnetic Resonance Imaging in the Pre-Treatment Staging of Prostate Cancer Guideline Development Group

An assessment conducted in January 2023 placed Guideline 27-3 IN REVIEW. This means that it is undergoing a review for currency and relevance. It is still appropriate for this document to be available while this updating process unfolds. The PEBC has a formal and standardized process to ensure the currency of each document (PEBC Assessment & Review Protocol)

Guideline 27-3 is comprised of 5 five sections. You can access the summary and full report here:
https://www.cancercareontario.ca/en/guidelines-advice/types-of-cancer/606

Report Date: March 29, 2016

For information about this document, please contact M Haider,
the corresponding author, through the PEBC via:
Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

For information about the PEBC and the most current version of all reports, please visit the CCO website at http://www.cancercare.on.ca/ or contact the PEBC office at:
Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

PEBC Report Citation (Vancouver Style): Haider M, Salerno J, Finelli A, Morash C, Morgan S, Power N, Schieda N; Magnetic Resonance Imaging in the Pre-Treatment Staging of Prostate Cancer Guideline Development Group. Magnetic resonance imaging in staging for prostate cancer. Toronto (ON): Cancer Care Ontario; 2016 Mar 29 [In Review 2023 Jan]. Program in Evidence-based Care Guideline No.: 27-3 IN REVIEW.

Journal Citation (Vancouver Style): Salerno J, Finelli A, Morash C, Morgan SC, Power N, Schieda N, Haider MA. Multiparametric magnetic resonance imaging for pre-treatment local staging of prostate cancer: A Cancer Care Ontario clinical practice guideline. Can Urol Assoc J. 2016 Sep-Oct;10(9-10):E332-E339.

Conference Citation (Vancouver Style): Fergenbaum J, Haider M, Finelli T, Morash C, Morgan S, Power N, et al. Development of a clinical practice guideline on magnetic resonance imaging of prostate cancer [abstract]. Presented at: American College of Epidemiology (ACE) 2015 Annual Meeting; 2015 Sep 26-29; Atlanta, Georgia, USA.

Table of Contents

Magnetic Resonance Imaging for Pre-Treatment Local Staging of Prostate Cancer

Section 1: Recommendations

This section is a quick reference guide and provides the guideline recommendations only. For key evidence associated with each recommendation, see Section 2.

GUIDELINE OBJECTIVES

To make recommendations with respect to the use of T2-weighted magnetic resonance imaging (MRI) ± functional sequences in the pre-treatment local staging of patients with newly diagnosed prostate cancer.

Note:

- MRI refers to T2-weighted MRI.
- Functional sequences include dynamic contrast-enhanced imaging (DCE), diffusion-weighted imaging (DWI), and proton magnetic resonance spectroscopic imaging (MRS).
- In this guideline, the terminology of MRI ± functional sequences is used interchangeably with MRI ± DCE, DWI, and MRS (See Glossary of Terms, Appendix 1).

TARGET POPULATION

Men with newly diagnosed biopsy-confirmed prostate cancer who are under consideration for radical treatment.

INTENDED USERS

Clinicians who are involved in the staging and treatment of prostate cancer patients.

RECOMMENDATIONS

Recommendation 1

Multiparametric MRI (mpMRI) use for pre-treatment local staging of prostate cancer is a reasonable option for assessment of extraprostatic extension (EPE) in intermediate- and high-risk patients being considered for radical therapy if knowledge of EPE will alter management.

Qualifying Statements for Recommendation 1

- mpMRI is the addition of two or more functional sequences to T2-weighted MRI such as DCE, DWI, and MRS imaging. See Appendix 10 for mpMRI and technical specifications [1].
- Prostate cancer risk groups are defined according to the Prostate Cancer Treatment Pathway [2]. See Appendix 2 for risk level definitions.
- Based on consensus expert clinical opinion, pre-treatment local staging of selected newly diagnosed intermediate-risk prostate cancer patients by mpMRI could be beneficial as it would add useful information regarding the characteristics of the intraprostatic tumour such as its location and extent, as well as inform on the presence of EPE to help aid in treatment planning.
- Due to the relatively low sensitivity of pre-treatment local staging by MRI (± functional sequences) for the detection of EPE and, thus, the possibility of false-negative findings, the clinical management of those patients receiving negative results should be considered on an individual basis in the context of pre-treatment

clinical nomograms. This particularly applies to high-risk surgical patients. Caution should be exercised when considering nerve-sparing surgery on the basis of mpMRI evaluation indicating no EPE on the side of prostate cancer.

Recommendation 2
Centres using mpMRI for local prostate cancer staging must have a quality assurance program in place to measure diagnostic performance of mpMRI.

Qualifying Statements for Recommendation 2

- This recommendation is based on consensus expert clinical opinion. Use of standardized reporting has shown statistically significant improvements in sensitivity [3] warranting further consideration of mpMRI quality assurance programs.
- mpMRI is the addition of two or more functional sequences to T2-weighted MRI such as DCE, DWI, and MRS imaging. See Appendix 10 for mpMRI and technical specifications [1].

Magnetic Resonance Imaging for Pre-Treatment Local Staging of Prostate Cancer

Section 2: Guideline – Recommendations and Key Evidence

GUIDELINE OBJECTIVES

To make recommendations with respect to the use of T2-weighted magnetic resonance imaging (MRI) ± functional sequences in the pre-treatment local staging of patients with newly diagnosed prostate cancer.

Note:

- MRI refers to T2-weighted MRI.
- Functional sequences include dynamic contrast-enhanced imaging (DCE), diffusion-weighted imaging (DWI), and proton magnetic resonance spectroscopic imaging (MRS).
- In this guideline, the terminology of MRI ± functional sequences is used interchangeably with MRI ± DCE, DWI, and MRS (See Glossary of Terms, Appendix 1).

TARGET POPULATION

Men with newly diagnosed biopsy-confirmed prostate cancer who are under consideration for radical treatment.

INTENDED USERS

Clinicians who are involved in the staging and treatment of prostate cancer patients.

RECOMMENDATIONS, KEY EVIDENCE, AND INTERPRETATION OF EVIDENCE

Recommendation 1
Multiparametric MRI (mpMRI) use for pre-treatment local staging of prostate cancer is a reasonable option for assessment of extraprostatic extension (EPE) in intermediate- and high-risk patients being considered for radical therapy if knowledge of EPE will alter management.
Qualifying Statements for Recommendation 1
• mpMRI is the addition of functional sequences to T2-weighted MRI such as DCE, DWI, and MRS imaging. See Appendix 10 for mpMRI and technical specifications [1]. • Prostate cancer risk groups are defined according to the Prostate Cancer Treatment Pathway [2]. See Appendix 2 for risk level definitions. • Based on consensus expert clinical opinion, pre-treatment local staging of selected newly diagnosed intermediate-risk prostate cancer patients by mpMRI could be beneficial as it would add useful information regarding the characteristics of the intraprostatic tumour such as its location and extent as well as inform on the presence of EPE to help aid in treatment planning. • Due to the relatively low sensitivity of pre-treatment local staging by MRI ± functional sequences for the detection of EPE and, thus, the possibility of false-negative findings, the clinical management of those patients receiving negative results should be considered on an individual basis in the context of pre-treatment clinical nomograms. This particularly applies to high-risk surgical patients. Caution should be exercised when considering nerve-sparing surgery on the basis of mpMRI evaluation indicating no EPE on the side of prostate cancer.
Key Evidence for Recommendation 1

- For the outcome of diagnostic accuracy, 49 of 61 (80.3%) studies reported on the sensitivity and specificity of MRI ± DCE, DWI, MRS [3-51]. There were 19 of 49 (38.8%) studies [4,6,12,13,15,19,20,23,25,27-29,31,32,34,36,37,39,51] of ≥1.5 T MRI ± DCE, DWI, MRS included in scatterplot analysis of sensitivity versus 1 – specificity that considered the detection of EPE and/or seminal vesicle invasion (SVI) with the use of an endorectal coil (ER) (EPE in 16 studies [6,12,13,19,20,23,25,27-29,31,34,36,37,39,51] and SVI in 13 studies [4,12,13,15,20,23,25,27,32,34,36,37,39]).
- For EPE and SVI, the display of data showed a range of performance and diagnostic accuracies, between 14.0% and 97.0% in terms of sensitivity and between 74.0% and 100% in terms of specificity (19 studies). Notably, the range of values for sensitivity in the detection of EPE was wider (14.0% to 90.0%) (16 studies) compared with the range of values for sensitivity in the detection of SVI (34.9% to 97.0%) (13 studies).
- Considering ER status, compared with studies that used an ER, lower sensitivities were shown for studies of 1.5 T MRI ± DCE, DWI, MRS that did not use an ER (range: 0% to 81.3%) [7,18,21,22,24,30,35,39,41] (nine studies), whereas sensitivity was maintained for studies of 3 T MRI ± DCE, DWI, MRS that did not use an ER (range: 22.0% to 92.0%) [5,8,9,11,14,16,26,33,36] (13 studies).
- In summary, the median sensitivities (SN) and specificities (SP) in the pre-treatment local staging of prostate cancer are as follows:
    - SN: 50.0% and SP: 91.0% for 1.5 T + 3 T MRI + ER ± DCE, DWI, MRS in the detection of EPE (16 studies).
    - SN: 50.0% and SP: 96.0% for 1.5 T + 3 T MRI + ER ± DCE, DWI, MRS in the detection of SVI (13 studies).
    - SN: 36.2% and SP: 90.3% for 1.5 T MRI ± DCE, DWI, MRS without an ER (nine studies).
    - SN: 58.3% and SP: 86.6% for 3 T MRI ± DCE, DWI, MRS without an ER (13 studies).
- Based on an overall rating of moderate quality, a recent meta-analysis examined the diagnostic accuracy of MRI ± DCE, DWI, MRS and showed the sensitivity and specificity for EPE as 0.57 (95% confidence interval [CI], 0.49 to 0.64) and 0.91 (95% CI, 0.88 to 0.93) and for SVI as 0.58 (95% CI, 0.47 to 0.68) and 0.96 (95% CI, 0.95 to 0.97), respectively [52].
- There were six of 61 (9.8%) studies that reported on the outcome of change in treatment plan [5,14,41,45,51,53]. All six studies were consistent in showing increased therapy, with <1% to 43% of patients experiencing increased therapy due to staging on MRI ± DCE, DWI, MRS [5,14,41,45,51,53]. Three studies reported that MRI ± DCE, DWI, MRS-based treatment plans were correct, as shown in 63% to 97% of patients [14,51,53].
- Twenty-one of 61 (34.4%) studies reported on the outcome of change in stage classification. There were 20 of 21 (95.2%) studies that consistently demonstrated upstaging [41], and upstaging as compared with routine clinical staging [5,6,12,18-21,24,27,28,30,31,35,45,53-57]. Staging by MRI ± DCE, DWI, MRS was correct by pathology in seven studies, ranging from 11% to 85% [19-21,35,45,53,56].

Interpretation of Evidence for Recommendation 1

- The results of the current synthesis of recently published primary studies on diagnostic accuracy are consistent with a published meta-analysis and systematic review. Both the current systematic review and the recently published meta-analysis show modest sensitivities (e.g., 50% to 60%) and excellent specificities (e.g., >85%),

suggesting the robustness of the evidence base and the beneficial performance of MRI (1.5 T + ER and 3 T ± ER) ± DCE, DWI, MRS.

- Based on consensus expert clinical opinion, above-average sensitivities were achieved in some studies as shown by the display of data. Those studies tended to comprise intermediate-risk patients. Quality assessment revealed a high risk of bias for diagnostic accuracy outcomes. Improvements in study design in terms of patient sampling and study design, larger studies, standardized use of functional sequences, explicit pathology criteria, and blinding of both radiologists and pathologists are needed to improve the quality of the evidence.
- Beneficial effects of imaging were shown for the outcomes of treatment plan and staging classification. Quality assessment revealed serious risk of bias for the non-diagnostic accuracy outcomes. The limitations of the studies included the small study sizes, the paucity of clinical and patient outcome data, the lack of consistently reported outcomes across studies, and the lack of comparable analysis methods across studies, making it difficult to draw conclusions regarding the impact of MRI ± DCE, DWI, MRS on clinical and patient outcomes.
- Although beneficial effects of imaging have been shown, a recent randomized controlled trial did not show a beneficial effect of MRI+DWI [41]. There were a number of limitations to the trial that diluted the ability to detect a difference in surgical margin status between MRI+DWI and non-MRI groups including: limited power; specified criteria for deciding how to modify the surgical plan based on imaging including a wider excision at sites of tumour was not part of the study design; limitations of the surgical technique associated with robotic surgery including the lack of a specified surgical protocol for various types of imaging findings; and the protocol detailing communication and patient handling between the radiologist and the urologist could be improved.

Recommendation 2
Centres using mpMRI for local prostate cancer staging must have a quality assurance program in place to measure diagnostic performance of mpMRI.
Qualifying Statements for Recommendation 2
• This recommendation is based on consensus expert clinical opinion. Use of standardized reporting has shown statistically significant improvements in sensitivity [3] warranting further consideration of mpMRI quality assurance programs. • mpMRI is the addition of two or more functional sequences to T2-weighted MRI such as DCE, DWI, and MRS imaging. See Appendix 10 for mpMRI and technical specifications [1].

FURTHER QUALIFYING STATEMENTS

Please see "MpMRI and Technical Specification" in Appendix 10.

IMPLEMENTATION CONSIDERATIONS

The following implementation themes may be considered: availability of scanners, clinician or health administration uptake, and having a local quality assurance program in place (e.g., including measuring performance and diagnostic accuracy, tracking outcomes, radiologist training/education/experience, and ongoing peer review process).

RELATED GUIDELINES

- #27-2 Multiparametric Magnetic Resonance Imaging in the Diagnosis of Clinically Significant Prostate Cancer
- #17-9 Active Surveillance for the Management of Localized Prostate Cancer

Magnetic Resonance Imaging for Pre-Treatment Local Staging of Prostate Cancer:

Section 3: Guideline Methods Overview

This section summarizes the methods used to create the guideline. For the systematic review, see Section 4.

THE PROGRAM IN EVIDENCE-BASED CARE

The Program in Evidence-Based Care (PEBC) is an initiative of the Ontario provincial cancer system, Cancer Care Ontario (CCO). The PEBC mandate is to improve the lives of Ontarians affected by cancer through the development, dissemination, and evaluation of evidence-based products designed to facilitate clinical, planning, and policy decisions about cancer control.

The PEBC supports the work of Guideline Development Groups (GDGs) in the development of various PEBC products. The GDGs are composed of clinicians, other healthcare providers and decision makers, methodologists, and community representatives from across the province.

The PEBC is a provincial initiative of CCO supported by the Ontario Ministry of Health and Long-Term Care (OMHLTC). All work produced by the PEBC is editorially independent from the OMHLTC.

BACKGROUND FOR GUIDELINE

The Cancer Imaging Program of CCO in collaboration with the Prostate Cancer Disease Pathway Management Secretariat and the Genitourinary Group identified a need to evaluate T2-weighted magnetic resonance imaging (MRI) in the pre-treatment local staging of prostate cancer. The technology is currently in use with the potential for practice variation and inappropriate use (under- or overuse); there exists the potential for a system-wide impact with its ongoing use; and is currently articulated for use within the Prostate Cancer Treatment Pathway, although guidance around its use is lacking. Renewed interest in MRI has also mounted due to the increased awareness of the utility of multiparametric MRI (mpMRI). MpMRI is the addition of functional sequences to MRI such as dynamic contrast-enhanced imaging (DCE), diffusion-weighted imaging (DWI), and proton magnetic resonance spectroscopic imaging (MRS). Current modalities of pre-treatment local staging of prostate cancer have drawbacks including under- or overestimating the extent of disease. Information about the stage of disease is used for surgical and treatment planning that is associated with potential downstream morbidity and mortality concerns for patients. It is not known which patients may benefit from MRI ± functional sequences including DCE, DWI, and MRS in the pre-treatment local staging of prostate cancer. Taken together, with the prevalence of prostate cancer and the value in cancer localization by MRI ± DCE, DWI, and MRS, it was sought to produce guidance on MRI ± functional sequences in the pre-treatment local staging of prostate cancer.

GUIDELINE DEVELOPERS

This guideline was developed by the Magnetic Resonance Imaging in the Pre-Treatment Staging of Prostate Cancer GDG (Appendix 3), which was convened at the request of the Cancer

Imaging Program of CCO in collaboration with the Prostate Cancer Disease Pathway Management Secretariat and the Genitourinary Group.

The project was led by a small Working Group of the Magnetic Resonance Imaging in the Pre-Treatment Staging of Prostate Cancer GDG, which was responsible for reviewing the evidence base, drafting the guideline recommendations and responding to comments received during the document review process. The Working Group had expertise in radiology, urology, radiation oncology, and health research methodology. Other members of the Magnetic Resonance Imaging in the Pre-Treatment Staging of Prostate Cancer GDG served as the Expert Panel and were responsible for the review and approval of the draft document produced by the Working Group. Conflict of interest declarations for all GDG members are summarized in Appendix 3, and were managed in accordance with the PEBC Conflict of Interest Policy.

GUIDELINE DEVELOPMENT METHODS

The PEBC produces evidence-based and evidence-informed guidance documents using the methods of the Practice Guidelines Development Cycle [58]. This process includes a systematic review, interpretation of the evidence by the Working Group and draft recommendations, internal review by content and methodology experts, and external review by Ontario clinicians and other stakeholders.

The PEBC uses the AGREE II framework [59] as a methodological strategy for guideline development. AGREE II is a 23-item validated tool that is designed to assess the methodological rigour and transparency of guideline development.

The currency of each document is ensured through periodic review and evaluation of the scientific literature and, where appropriate, the addition of newer literature to the original evidence base. This is described in the *PEBC Document Assessment and Review Protocol*. PEBC guideline recommendations are based on clinical evidence, and not on feasibility of implementation; however, a list of implementation considerations such as costs, human resources, and unique requirements for special or disadvantaged populations is provided along with the recommendations for information purposes. PEBC guideline development methods are described in more detail in the *PEBC Handbook* and the *PEBC Methods Handbook*.

Search for Existing Guidelines

A search for existing guidelines is generally undertaken prior to searching for existing systematic reviews or primary literature. This is done with the goal of identifying existing guidelines for adaptation or endorsement in order to avoid the duplication of guideline development efforts across jurisdictions. For this project, the following sources were searched for existing guidelines that addressed the research questions:

- Practice guideline databases: Standards and Guidelines Evidence (SAGE), National Guidelines Clearinghouse, and the Canadian Medical Association Journal Infobase
- Guideline developer websites and other international/urological associations:
    - National Institute of Clinical Excellence
    - Scottish Intercollegiate Guidelines Network
    - American Society of Clinical Oncology
    - Australian National Health and Medical Research Council
    - New Zealand Guidelines Group
    - American Urological Association
    - European Association of Urology
    - Canadian Urological Association
    - American College of Radiology
    - European Society of Urogenital Radiology
    - National Comprehensive Cancer Network

    - Chinese Urological Association
    - Singapore Urological Association
    - Taiwan Cooperation Oncology Group
    - U.S. Preventive Services Task Force
- MEDLINE and EMBASE databases: a systematic literature search for guidelines was performed.

The following criteria were used to select potentially relevant guidelines: publication year of 2005 to 2015 (past 10 years), guideline methods were well-described, and recommendations were articulated. Using MEDLINE and EMBASE databases, a systematic literature search for existing guidelines was conducted. If appropriate, guidelines that were considered relevant to the objectives and the research questions were then evaluated for quality using the AGREE II instrument. However, a search for existing guidelines for adaptation or endorsement did not yield an appropriate source document. Therefore, the AGREE II instrument was not used and a search of the primary literature was required (see Section 4).

GUIDELINE REVIEW AND APPROVAL

Internal Review

For the guideline document to be approved, 75% of the content experts who comprise the GDG Expert Panel must cast a vote indicating whether or not they approve the document, or abstain from voting for a specified reason, and of those that vote, 75% must approve the document. In addition, the PEBC Report Approval Panel (RAP), a three-person panel with methodology expertise, must unanimously approve the document. The Expert Panel and RAP members may specify that approval is conditional, and that changes to the document are required. If substantial changes are subsequently made to the recommendations during external review, then the revised draft must be resubmitted for approval by RAP and the GDG Expert Panel.

External Review

Feedback on the approved draft guideline is obtained from content experts and the target users through two processes. Through the Targeted Peer Review, several individuals with content expertise are identified by the GDG and asked to review and provide feedback on the guideline document. Through Professional Consultation, relevant care providers and other potential users of the guideline are contacted and asked to provide feedback on the guideline recommendations through a brief online survey. This consultation is intended to facilitate the dissemination of the final guidance report to Ontario practitioners.

ACKNOWLEDGEMENTS

The Magnetic Resonance Imaging in the Pre-Treatment Staging of Prostate Cancer GDG would like to thank the following individuals for their assistance in developing this report:

- Melissa Brouwers, Sheila McNair, Hans Messersmith, Erin Kennedy, Roxanne Cosby, RAP Reviewers, and Targeted Peer Reviewers for providing feedback on draft versions.
- Jimmy Zhang for conducting a data audit.
- Sara Miller for copy editing.

Magnetic Resonance Imaging for Pre-Treatment Local Staging of Prostate Cancer

Section 4: Systematic Review

INTRODUCTION

Prostate cancer among men in Canada ranks first in terms of the number of new cases of cancer (24%) and third in terms of mortality (10.1%), after colorectal cancer (12.4%) and lung cancer (26.6%) [60].

Controversy still exists in the care and management of prostate cancer. Ways in which prostate cancer is conventionally detected includes prostate-specific antigen (PSA) screening, digital rectal examination (DRE), and biopsy-confirmed diagnosis by transrectal ultrasonography (TRUS). Staging of prostate cancer helps to determine the extent and location of disease, with knowledge of the severity of disease from staging used to determine prognosis and inform treatment planning. Limitations when using traditional modalities to locally stage prostate cancer for treatment planning purposes include low specificity (e.g., screening), missed regions of the prostate during clinical examination or biopsy, and limited information on volume, extent, and aggressiveness of disease (e.g., DRE, TRUS). T2-weighted magnetic resonance imaging (MRI) with or without at least one functional sequence (± functional sequences) may better depict the zonal anatomy, pathology, and functionality of the prostate compared with conventional methods, thereby improving local staging and optimal treatment planning for prostate cancer patients [61].

MRI is a noninvasive tool to examine the anatomy and pathology of prostate cancers. Conventional localization of prostate cancer by MRI involves T1/T2-weighted reconstructed three-dimensional images that require radiologist interpretation. The magnetic field strength may be 1.5 Tesla (T) or 3 T, with the presence or absence of an endorectal coil (ER); however, the use of an ER is generally considered to provide optimal MRI quality [62]. Multiparametric MRI (mpMRI) offers a technical enhancement to conventional MRI by the addition of (one or more) functional sequences. Three functional sequences exist including dynamic contrast-enhanced imaging (DCE), diffusion-weighted imaging (DWI), and proton magnetic resonance spectroscopic imaging (MRS). DCE uses a contrast agent that allows perfusion over time to be displayed and interpreted. DWI measures the diffusion of water molecules through tissue, where prostate cancer has a reduced diffusion of water compared with normal tissue. Quantification of DWI using apparent diffusion coefficient maps can be performed. MRS reflects metabolite levels in the prostate such as choline, creatinine, and citrate [63].

Given the need to inform treatment planning across stages of prostate cancer beyond the information gathered on risk of spread by established nomograms, an examination of the use of MRI ± functional sequences in the pre-treatment local staging for patients with newly diagnosed prostate cancer was undertaken. In order to make recommendations as part of a clinical practice guideline, the Working Group of the Magnetic Resonance Imaging in the Pre-Treatment Staging of Prostate Cancer Guideline Development Group developed this evidentiary base upon which those recommendations are based. Based on the objectives of the guideline, the Working Group derived the research questions outlined below.

RESEARCH QUESTIONS

a) What is the performance and diagnostic accuracy of MRI ± DCE, DWI, MRS in men with newly diagnosed biopsy-confirmed prostate cancer who are under consideration for radical treatment?

b) What is the impact of pre-treatment local staging by MRI ± DCE, DWI, MRS on patient outcomes, biochemical recurrence, changes in treatment planning (including nerve-sparing surgery), changes in stage classification, and surgical margin status in men with newly diagnosed biopsy-confirmed prostate cancer who are under consideration for radical treatment?

Note:

- MRI refers to T2-weighted MRI.
- Functional sequences include DCE, DWI, and MRS.
- In this guideline, the terminology of MRI ± functional sequences is used interchangeably with MRI ± DCE, DWI, MRS (See Glossary of Terms, Appendix 1).

Additionally, in this guideline, radical treatment is defined as radical prostatectomy or external beam radiation therapy. Local staging is defined as staging of the prostate and the local tissues in the pelvis.

METHODS

This evidence review was conducted in two planned stages, including a search for systematic reviews followed by a search for primary literature. These stages are described in subsequent sections.

Search for Existing Systematic Reviews

A search was conducted for existing systematic reviews. Identified systematic reviews were evaluated based on their clinical content and relevance. Relevant systematic reviews were assessed using the 11-item Assessment of Multiple Systematic Reviews (AMSTAR) [64] tool to determine whether existing systematic reviews met a minimum threshold for methodological quality and could be considered for inclusion in the evidence base (Appendix 4). Systematic reviews published as a component of practice guidelines were also considered eligible for inclusion.

A priori, a systematic review as a component of a practice guideline was identified with a satisfactory AMSTAR score (score ≥7), which informed the 2013 starting search date for the literature search. Details of the literature search for systematic reviews are shown in Appendix 5.

In brief, it followed:

- Databases searched: MEDLINE, EMBASE, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Health Technology Assessment
- January 1, 2013 to February 17, 2016
- Search terms: prostate cancer, staging, and MRI
- Studies that examined ≥1.5 T MRI ± DCE, DWI, MRS in the pre-treatment local staging of prostate cancer

Any identified reviews that did not meet the criteria above, whose AMSTAR assessments indicated important deficiencies in quality, or that were otherwise not incorporated as part of the evidence base, were reported in the results and reference list, but not further described or discussed in detail.

Search for Primary Literature

For this clinical practice guideline, the search for primary literature builds upon previous work. A comprehensive and current systematic literature search was conducted in the relevant databases for the years of 2013 to 2016 (Appendix 4). The evidentiary base covering prior years (2008 to 2013) was comprised of studies already identified from a previous relevant systematic review that was identified and for which the objectives, questions, and outcomes were met for the purpose of this clinical practice guideline.

Literature Search Strategy

A literature search was conducted on February 17, 2016 using MEDLINE and EMBASE and other databases. Details of the literature search (main search and update) can be found in Appendix 5.

Study Selection Criteria and Process

Inclusion Criteria

- Studies published between January 1, 2008 and February 17, 2016
- English language, humans, adults ≥18 years of age
- Studies on pre-treatment local staging of prostate cancer
- Studies of men with newly diagnosed biopsy-confirmed prostate cancer who are candidates for radical treatment
- Studies using MRI of at least 1.5 T or higher ± ER ± DCE, DWI, MRS
- Studies that are systematic reviews, meta-analyses, randomized controlled trials, prospective or retrospective observational studies that answer the research questions of interest
- Studies reporting at least one outcome of interest
- Studies of a minimum size of 30 patients

Exclusion Criteria

- Case reports (n=1), conference abstracts, in vitro, or animal studies
- Studies of the technical aspects of MRI ± DCE, DWI, MRS
- Studies of prostate cancer patients post-treatment (e.g., hormonal or focal therapy or surgery)
- Studies pre-biopsy or pre-diagnosis/detection of prostate cancer
- Studies of combined technologies (e.g., MRI-guided TRUS biopsy in the diagnosis of prostate cancer, positron emission tomography)
- Studies in which patients have been previously diagnosed and clinically staged and are under active surveillance
- Studies in which the study methods are not well-described or not clear

A review of the titles and abstracts that resulted from the search was conducted by one reviewer (JS). For those items that warranted full-text review, one reviewer reviewed each item (JS).

Data Extraction and Assessment of Study Quality and Potential for Bias

Data abstraction was performed by one abstractor (JS). Abstracted data included study variables such as author, publication year, study location, and type of study; detailed study characteristics such as sample size, method of patient enrollment, inclusion/exclusion criteria, baseline staging method, presence of biopsy-proven prostate cancer, number of cores, time to imaging (e.g., median), median or mean age, median or mean PSA level (low: <10 ng/mL, intermediate: 10-20 ng/mL, high: >20 ng/mL), and pathological stage by radical prostatectomy

(gold standard/reference); imaging (index test) variables such as whether an experienced reader/radiologist performed image interpretation, whether the reader/radiologist was partially or completed blinded to clinical and/or pathology data, magnetic field strength, presence/absence of the use of an ER and/or body/pelvic coil, and whether one or more functional sequences were used including DCE, DWI, and MRS. Details of the definitions of positivity for tumour staging including extraprostatic extension (EPE) and seminal vesicle invasion (SVI) by each imaging platform were also abstracted. Additional information included the type of radical prostatectomy surgery (open, laparoscopic, robotic-assisted) and whether local tumour staging determined on imaging was used to inform the nature of radical prostatectomy surgery.

The following outcome information was abstracted from included studies: diagnostic accuracy (sensitivity, specificity, positive predictive value, and negative predictive value); clinical outcomes such as tumor, node, metastasis (TNM) stage classification, risk stratification category, treatment plan, positive surgical margin, biochemical recurrence; and patient outcomes. All outcomes were deemed primary outcomes.

A study was classified as having used an 'experienced' reader/radiologist when any level of experience was reported and as having the reader/radiologist as 'complete/partial blinded' when any evidence of blinding was reported.

Studies were categorized as having a retrospective study design if the imaging and outcome information had been collected in the past and prior to the start of the current study (e.g., historical medical record review study, a study that reviewed a prospectively maintained database). A study was categorized as having a prospective study design if the imaging information was collected at baseline; however, the outcome information for patients was collected at a point in time beyond the starting point of the study (e.g., biochemical recurrence at follow-up). Studies could display features of both retrospective and prospective designs. The terminology of case series was used in place of single-arm cohort study.

For data abstracted on baseline staging, if the method of clinical staging was not explicitly stated or was unclear but a biopsy Gleason score was reported in the text or tables, then it was assumed that clinical staging was performed by biopsy. Similarly, if biopsy-proven cancer was not explicitly stated but a biopsy Gleason score was reported in the text or tables, then it was assumed that study patients had prostate cancer that was biopsy-confirmed. Furthermore, if more than one clinical staging method was reported in the original paper without clarity around one particular method that was used for baseline staging, then all information was abstracted and reported (e.g., DRE, PSA, and biopsy). If baseline staging was reported to have been performed by clinical staging without further specifying details, then clinical staging was abstracted and reported as the baseline staging method.

Details of the imaging technology used were particularly noted when abstracting the information on diagnostic accuracy. If MRI plus one or more of DCE, DWI, MRS were used, then it was assumed (if not stated clearly) that the outcome measure reflects the simultaneous evaluation and interpretation of tumour stage by use of MRI plus one or more of DCE, DWI, MRS, and this was reported (e.g., MRI+DCE). However, if each functional sequence and its corresponding measure of diagnostic accuracy were reported individually and separately from MRI, then each imaging technology and its corresponding outcome information was reported (e.g., MRI, DCE, DWI, MRS). Therefore, an individual study could contribute more than one data point. All extracted data and information were audited by an independent auditor.

Heterogeneity

A priori, the variables including field strength, ER use, functional sequences, PSA level, and sample size were identified as potential sources of heterogeneity.

Risk of Bias

Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [65] was used to evaluate the risk of bias for the outcome of diagnostic accuracy. The Cochrane Risk of Bias Assessment Tool for Non-Randomized Studies of Interventions (ACROBAT-NRSI) was used to assess the risk of bias for the non-diagnostic accuracy outcomes [66].

Synthesizing the Evidence

All analysis included studies of MRI ± DCE, DWI, MRS. For studies with suitable data, the following outcome information was shown in descriptive tables including diagnostic accuracy (sensitivity, specificity, positive predictive value, and negative predictive value); clinical outcomes including TNM stage classification, risk stratification category, treatment plan, positive surgical margin, biochemical recurrence; and patient outcomes. Studies that reported other indices of diagnostic accuracy besides those mentioned above were summarized separately. Change in TNM stage was calculated from baseline clinical stage and stage at follow-up post-imaging. Change in treatment plan was calculated from the pre-imaging treatment plan and the post-imaging treatment plan. The percentage of positive surgical margins was reported separately for those studies that indicated that the imaging data were used to inform the nature of radical prostatectomy surgery. For outcomes identified in studies that did not provide data for a quantitative evidence synthesis, then a qualitative narrative summary was performed.

For the diagnostic accuracy outcomes of sensitivity and specificity among studies with paired data, the median and range of values were calculated and examined. Scatterplots of sensitivity versus 1 – specificity were generated and examined based on published data. The display of data was examined for potential outliers, and if detected, these studies were then further excluded from the analysis. There were two subgroups of interest identified a priori: (1) studies of 1.5 T MRI + ER ± DCE, DWI, MRS; and (2) studies of 3 T MRI + ER ± DCE, DWI, MRS. Detection of EPE and SVI were analyzed for the two groups. Data for EPE also represent data from individual studies specifying extracapsular extension. Diagnostic accuracy was interpreted according to standard methods [67].

RESULTS

Search for Existing Systematic Reviews

A number of systematic reviews were identified from the literature search. However, due to their narrow focus [68], narrow search dates [69-72], not meeting the inclusion/exclusion criteria [73], or a combination [69-72], they were excluded. One systematic review published as a component of a clinical practice guideline [74] was relevant to the objectives, questions, and outcomes of this guideline and also had a satisfactory AMSTAR score (Appendix 4). This systematic review helped to inform the literature search strategy and evidentiary base (see above, Search for Primary Literature). A second meta-analysis and systematic review was identified that examined the outcome of diagnostic accuracy (sensitivity, specificity) [52] and also had a satisfactory AMSTAR score (Appendix 4). The meta-analysis results of this study will be discussed under the relevant outcome heading, Diagnostic Accuracy (see below). Therefore, there was one systematic review/meta-analysis included.

Search for Primary Literature

Literature Search Results

From 2013 to 2015, there were a total of 1989 citations (after duplicates removed) identified. Including all sources, there were 172 papers that underwent full-text review. After an updated literature search for the years 2015 to 2016, there were 62 studies that were included (Appendix 6).

Study Design

As shown in Appendix 7, Table 1, there were 61 primary studies included [3-50,53-57,75-81], of which one (1.6%) study was a randomized controlled trial [41], 35 (57.3%) studies were retrospective case series [3,5-9,13-16,19,20,24-27,29,30,32,34,36,43,45-47,49,50,56,57,75-80], 19 (31.1%) studies were prospective case series [4,10-12,17,21,22,28,31,33,35,37,38,40,44,48,51,53,81], and six (9.8%) studies included aspects of both designs [18,23,39,42,54,55].

Study Populations and Demographics

Studies were predominately from Europe (24 studies, 39.3%) [5,11-13,15,18,19,21,22,24,27,29-31,41,43-46,49,53,54,57,81], Asia (17 studies, 27.9%) [7-10,14,16,23,26,33,35,36,39,50,55,75,78,79], and the United States (15 studies, 24.6%) [6,20,25,28,32,34,38,40,42,47,51,56,76,77,80], followed by one (1.6%) study each from Canada [3], Australia [17], Brazil [37], and two (3.3%) studies from Iran [4,48] (Appendix 7, Table 1).

The patient populations across studies according to PSA were predominately low (41 studies, 67.2%) [3,5-11,14-16,20,24,25,28-30,33,34,37,38,40-43,45-47,49,51,53-57,75-78,80,81] followed by intermediate (19 studies, (31.1%) studies) [4,12,13,18,19,21-23,26,27,31,32,35,36,39,44,48,50,79]. No studies included patients with average population levels of PSA >20 ng/mL. There was one study (1.6%) that did not report the PSA level of its patients [17]. The size of studies ranged from studies as small as 31 patients [42] up to larger studies of 922 patients [23]. All studies showed evidence that patients had biopsy-confirmed prostate cancer and all studies had some form of clinical staging (Appendix 7, Table 2).

Imaging Characteristics

The use of MRI ± DCE, DWI, MRS is shown in Appendix 7, Table 3. Field strength of 1.5 T was used in 30 (49.1%) studies [4,7,10,12,15,18,21,22,24,25,27-30,32,35,37,39,41,46,48,49,51,54,56,57,75-77,80], whereas a field strength of 3 T was used in 30 (49.1%) studies [3,5,6,8,9,11,13,14,16,17,19,20,26,31,33,34,36,38,40,42-45,47,50,53,55,78,79,81]. The study by Jeong et al (2013) [23] included the use of both field strengths; however, the study was classified as 3 T MRI for analysis purposes since a larger proportion of patients underwent imaging at the higher field strength (3 T: 73.3% versus 1.5 T: 26.7%). There were 34 (55.7%) studies that used an ER, with or without body or pelvic coils [4,6,10,12,13,15,17,19,20,23,25,27-29,31,32,34,36-40,43,46,48,49,51,53,54,56,57,76,77,80]; however, there were only seven (11.5%) studies that used an ER alone [17,19,25,46,48,51,54]. There were five (8.2%) studies in which neither an ER nor body or pelvic coils were used [24,26,35,45,50]. More studies used an ER than not (56% versus 44%). The interpretation of images was most often performed by an experienced reader or radiologist (88.0%); however, a wide range in experience level was reported and considered. More than one-half (55.7%) of personnel were completely blinded or partially blinded including clinical or patient data with or without explicitly stating also being blinded to pathology data (reference standard) when interpreting MRI (index test). There were 17 (27.9%) studies that did not use any functional sequence [4,15,18,24,25,27-29,36,39,54,56,57,76-79]. If one functional sequence was used alone with conventional MRI, it was mostly DWI, as shown in 10 studies (16.4%) [7,8,11,16,21,23,26,41,50,75], followed by two (3.3%) studies that used MRI+DCE [30,35] and two studies (3.3%) that used MRI+MRS [48,80]. Twenty-one studies (34.4%) used MRI+DCE+DWI

[5,6,9,12,14,19,20,22,31-33,55], whereas two (3.3%) studies used MRI+DCE+MRS [37,40], and only one (1.6%) study used MRI+DWI+MRS [10]. MRI+DCE+DWI+MRS was used in six studies (9.8%) [13,17,34,38,51,53]. A summary of the descriptive characteristics of included primary studies is shown in Table 4-1.

Quality Assessment

Diagnostic Accuracy

The QUADAS-2 risk of bias tool was used to assess study quality for the 19 studies that contributed quantitative paired data to the outcome of diagnostic accuracy in terms sensitivity and specificity (Appendix 9, Tables 1 and 2).

For patient selection, each study included biopsy-confirmed prostate cancer patients; therefore, the risk of bias due to ill-defined disease status is low. Patients were individuals that would typically be scheduled to undergo a radical prostatectomy. The main risk of bias stems from the study design, where not all studies enrolled patients consecutively or used a random sample of eligible patients, suggesting the potential for selection bias. Overall, patient selection as a source of bias was judged to be mixed (low-high risk).

For the index test, there were 11 of 19 studies (57.9%) in which the radiologist interpreting the MRI images was reported not to have been blinded to pathology staging information during evaluation, which may lead to a high risk of bias. However, when blinding was reported, whether the presence of blinding referred to both clinical and pathology staging information was not always clear from the methods reported. There were 10 studies of 19 studies (52.6%) that used one or more functional sequences and variability in use to evaluate and define EPE and SVI may have introduced bias (e.g., lack of standardization, criteria or threshold not pre-specified). Overall, the index test as a source of bias was judged to be mixed (low-high risk).

For the reference standard, the gold standard of pathology review of radical prostatectomy specimens to determine histological grade was used for all studies, suggesting a low risk of bias. Although a number of studies did not report the time interval between MRI imaging and radical prostatectomy, it is unlikely that there was disease progression bias since prostate cancer is a slow-growing tumour. Not many studies reported that the pathologist was blinded to MRI data (index test); therefore, information bias cannot be ruled out. The pathology criteria used to evaluate EPE and SVI was not consistently reported; therefore, patients could have been misclassified based on pathology review. Overall, the reference standard (pathology) as a source of bias was judged to be mixed (low-high risk and unclear risk).

Although there are no applicability concerns, in summary, there are a number of methodological concerns that would contribute to a potential high risk of bias according to QUADAS-2 criteria for the outcome of diagnostic accuracy (sensitivity, specificity).

Table 4-1. Summary of Included Primary Studies (n=61)

Characteristic	Range or Number of Studies (%)
Age (mean or median)	58 – 70 years
Minimum time to imaging post-biopsy	2 weeks
Experienced radiologist	54 (88.5)
Radiologist blinding	34 (55.7)
Field strength	
1.5 T	30 (49.1)
3 T	30 (49.1)
Both	1 (1.6)
Study type	
RCT	1 (1.6)
Retrospective	35 (57.3)
Prospective	19 (31.1)
Mixed	6 (9.8)
PSA level (mean or median)	
<10 ng/mL	41 (67.2)
10–20 ng/mL	19 (31.1)
>20 ng/mL	0 (0)
Not known	1 (1.6)
No ER use	27 (44.3)
Plus body or pelvic coil	22 (36.1)
Minus body or pelvic coil	5 (8.2)
ER use	34 (55.7)
Plus body or pelvic coil	27 (44.3)
Minus body or pelvic coil	7 (11.5)
Imaging	
MRI (no functional sequences)	17 (27.9)
MRI + DWI	10 (16.4)
MRI + DCE	2 (3.3)
MRI + MRS	2 (3.3)
MRI + DCE + DWI	21 (34.4)
MRI + DCE + MRS	2 (3.3)
MRI + DWI + MRIS	1 (1.6)
DCE + DWI + MRS (all three)	6 (9.8)

Abbreviations: DCE, dynamic contrast-enhanced; DWI, diffusion-weighted imaging; ER, endorectal coil; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; PSA, prostate-specific antigen; RCT, randomized controlled trial; T, Tesla.

Quality Assessment

Clinical and Patient Outcomes

By outcome, there were 21 studies on stage classification; two studies on risk stratification; six studies on treatment plan; 19 observational studies and one randomized controlled trial on surgical margins; five studies on biochemical recurrence; and no studies with patient outcomes.

ACROBAT-NRSI was used to assess the quality of studies that contributed to the assessment of clinical outcomes. The main limitations are the non-standardized use of functional sequences including a lack of explicitly stated criteria to assess EPE and SVI during radiology review, and a lack of complete blinding of the radiologist during evaluation and interpretation. Moreover, factors such physician-to-physician communication become more important as the outcome of interest becomes increasingly complex; for example, how was the imaging information used to alter treatment plan or surgery, which then impacts the outcomes of positive surgical margins and biochemical recurrence. These important details were not reported in the studies. Therefore, according to ACROBAT-NRSI criteria, there exists "bias in the measurement of interventions" and "bias in the measurement of outcomes". For observational studies, the ACROBAT-NRSI domain of confounding automatically signals a moderate level of bias (moderate risk of bias defined as "the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial"). When MRI ± DCE, DWI, MRS stage is compared with routine clinical stage for the outcome of TNM stage classification, in addition to heterogeneity in MRI ± DCE, DWI, MRS use and the clinical staging information used, there are additional patient clinical differences between the two groups being compared on factors other than the staging modality, which may have introduced unaccounted for bias. However, both staging modalities were conducted on the same group of patients within each study and change in TNM stage classification was calculated within each stage category. Taken together, there are non-trivial serious risks of biases for the examined clinical outcomes (serious risk of bias defined as studies having some important problems).

OUTCOMES

Diagnostic Accuracy

Primary Studies

There were 52 of 61 (85.2%) studies that had information on diagnostic accuracy, and 49 of 61 (80.3%) studies had quantitative information on one or more measure of diagnostic accuracy including sensitivity, specificity, positive predictive value, and negative predictive value [3-51] (Appendix 7, Table 4). All studies used histological review by pathology of radical prostatectomy specimens from surgery as the gold standard/reference for staging. There were four studies that provided information on other measures of diagnostic accuracy (i.e., area under the curve) (Appendix 7, Table 5) [77,78,80,81]. Twenty-four of 49 studies (49.0%) did not use an ER and, therefore, were not included in the scatterplot analysis [3,5,7-9,11,14,16,18,21,22,24,26,30,33,35,36,39,41,42,44,45,47,50].

Studies Using an ER

When the results of the scatterplots for sensitivity versus 1 – specificity were examined by the a priori defined subgroups, the display of data for studies that used 1.5 T MRI+ER (± functional sequences) in the detection of EPE was similar to the display of data for studies that used 3 T MRI+ER (± functional sequences) in the detection of EPE. A similar trend was observed

for the detection of SVI. Consequently, the results for 1.5 T and 3 T MRI were combined (data not shown).

When all eligible studies for ≥ 1.5 T + ER ± DCE, DWI, MRS in the detection of EPE and SVI were plotted, there were two studies that were outliers, as shown by the display of data [43,46]. Consequently, these two studies were excluded from the analysis.

EPE Detection

Figure 4-1 shows that 1.5 T or 3 T MRI + ER ± DCE, DWI, MRS in the detection of EPE is associated with sensitivities >60% and corresponding specificities >70% to 80% in approximately one-quarter of studies (four of 16 studies, 25.0%) [6,13,27,34]. In those studies, there was a range of functional sequences used including DCE imaging and DWI in one study [6], and all three functional sequences in two studies [13,34]. One study did not use any functional sequences [27]. Note that there were 17 data points including 16 studies analyzed for EPE [6,12,13,19,20,23,25,27-29,31,34,36,37,39,51]. Among the top four studies in terms of sensitivity [6,13,27,34], the individual study populations varied in the risk of spread according to PSA level with two of four studies including intermediate-risk populations [13,27] and the remaining two studies including low-risk populations [6,34]. The highest sensitivity (90%) was shown among intermediate-risk patients with the combined reported use of MRI and all three functional sequences. Yet, this same study of 37 patients showed a lower specificity value (approximately 70%) [13]. Among the four studies, the display of data showed three of four studies used 3 T MRI [6,13,34]. Overall, there were few studies and too few data points in the detection of EPE above 60% sensitivity, which would indicate a moderate level of sensitivity, to draw robust conclusions by potential heterogeneity factors (e.g., field strength, PSA level, functional sequence). Notably, these four studies were of a small size (<155 subjects).

As seen in Figure 4-1, two larger studies showed lower sensitivity values <50% yet maintained specificity >80% [23,29]. Upon further examination, in the study by Roethke et al (2013) [29], when the sensitivity of 1.5 T MRI (no functional sequences) plus ER use in the detection of EPE was examined by classifying patients by the D'Amico risk classification scheme, sensitivity did not improve, with a range of 20.0% for patients having PSA <10 ng/mL (low) and a Gleason score ≤6 to 47.6% for patients having a PSA ≥10 ng/mL (intermediate/high) and/or a Gleason score ≥7. In the study by Jeong et al (2013) [23], when the sensitivity of predominately 3 T MRI+DWI + ER and phased array coil use in the detection of EPE was examined by clinical variables, notably increased sensitivity values were shown with clinical stage T2c-T3 (58.7%) patients and patients having three high-risk factors (56.8%). For this study, both 1.5 T (26.7% use) and 3 T (73.3% use) MRI were used during the course of the study and although high-risk patients were sought, 51.9% were staged as T1c and 41.9% were determined to have organ-confined disease at pathology. Therefore, despite further interrogation by study authors within each study as described above, sensitivity values did not exceed a modest value (approximately 50% to 60%); however, sensitivity was improved with elevated-risk patients. Characteristics of diagnostic accuracy studies from Figure 4-1 are summarized in Table 4-2.

As shown in Table 4-3, the median sensitivity in the detection of EPE is 50.0%, with a range of values of 14.0% to 90.0%. The corresponding median specificity is 91.0%, with a range of values of 74.0% to 98.0%.

SVI Detection

Figure 4-2 shows that 1.5 T or 3 T MRI + ER ± DCE, DWI, MRS in the detection of SVI is associated with sensitivities >60% and corresponding specificities >80%, as shown in approximately one-half of the studies (six of 13 studies, 46.2%) [4,12,13,27,32,34]. In those studies, there was a range of functional sequences used including DCE and DWI in two studies [12,32], and all three functional sequences in two studies [13,34]. Two studies did not use any

functional sequences [4,27]. One study contributed data to both MRI use alone and MRI plus DCE and DWI use, as noted above [32]. Note that there were 15 data points including 13 studies for the analysis of SVI [4,12,13,15,20,23,25,27,32,34,36,37,39]. A majority of studies with >60% sensitivity included intermediate-risk patients and 1.5 T; however, there are too few studies and too few data points to draw robust conclusions regarding heterogeneity factors (e.g., field strength, PSA level, functional sequence). Again, the large study (n=922 patients) by Jeong et al (2013) [23] showed lower sensitivity (34.9%). Increases in sensitivity in the detection of SVI is shown in the original study when examined by clinical variables including clinical stage T2c-T3 (51.0%) and patients having three high-risk factors (42.3%). Characteristics of diagnostic accuracy studies from Figure 4-2 are summarized in Table 4-2.

As shown in Table 4-3, the median sensitivity in the detection of SVI is 50.0%, with a range of values of 34.9% to 97.0%. The corresponding median specificity is 96.0%, with a range of values of 83.1% to 100%.

Studies Without Using an ER

A post hoc analysis of studies that did not use an ER was performed to help inform the recommendations (Appendix 7, Figures 1 and 2).

1.5 T

Appendix 7, Figure 1 shows the paired analysis of sensitivity versus 1 – specificity for 1.5 T MRI ± DCE, DWI, MRS without an ER in the detection of tumour (2 studies), EPE (7 studies) and SVI (3 studies) combined. There were 18 data points including nine studies [7,18,21,22,24,30,35,39,41]. The display of data shows a downward shift towards lower sensitivity values. As shown in Table 4-3, the median sensitivity for 1.5 T, no ER, ± functional sequences in the detection of tumor, EPE and SVI is 36.2%, with a range of values of 0% to 81.3%.

3 T

Appendix 7, Figure 2 shows the paired analysis of sensitivity versus 1 – specificity for 3 T MRI ± DCE, DWI, MRS without an ER in the detection of tumour (2 studies), EPE (11 studies) and SVI (2 studies) combined. There were 20 data points including 13 studies [3,5,8,9,11,14,16,26,33,36,44,47,50]. The median sensitivity for 3 T, no ER, ± functional sequences in the detection of tumour, EPE and SVI is 58.3%, with a range of values of 22.0% to 92.0%.

Published Systematic Review/Meta-Analysis

A recent systematic review and meta-analysis examined the diagnostic accuracy of MRI with or without functional sequences for local staging in men with biopsy-confirmed prostate cancer using radical prostatectomy as the reference gold standard [52]. The literature search for studies was between the years of 2000 to August 12, 2014. The results of the literature search identified 75 studies, with 45 studies on EPE, 34 studies on SVI, and 38 studies on overall stage T3. Studies were included if the 2 × 2 tables for diagnostic accuracy could be reconstructed. The analysis involved pooling original data from included studies and contact with original authors for studies that did not report sufficient data. The sensitivity and specificity with corresponding 95% confidence intervals were calculated. The overall quality of evidence was judged by the authors to be moderate. The results of this meta-analysis are shown in Table 4-2. Overall, the results are similar to the analysis of published data in the current synthesis.

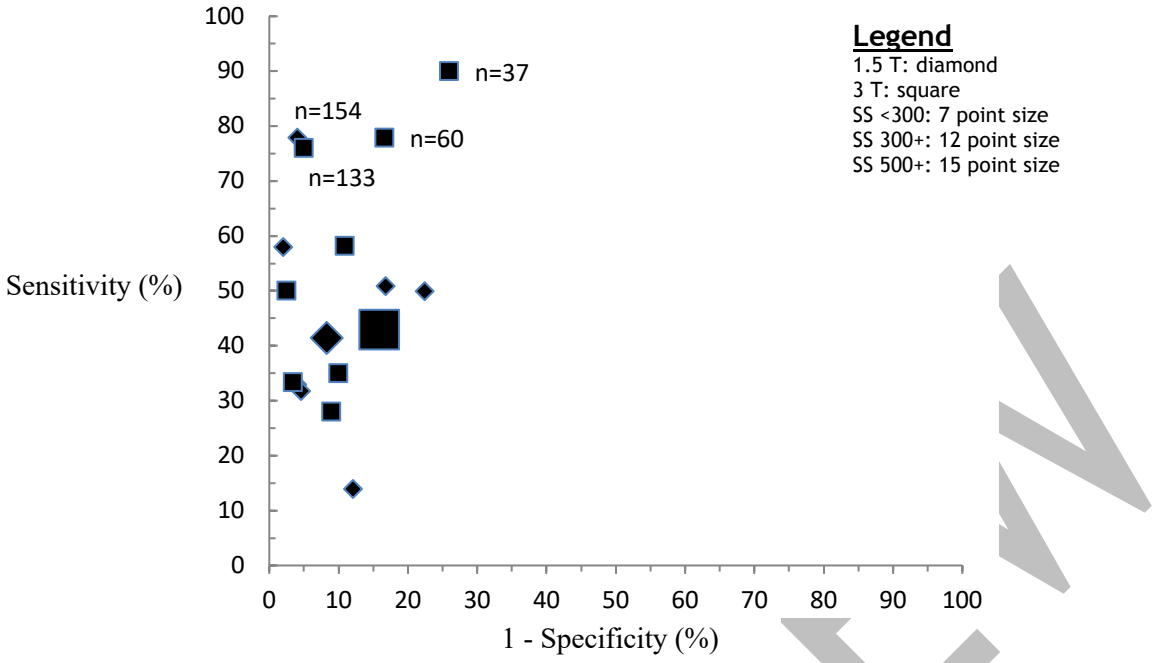


Figure 4-1. MRI Studies of 1.5 T + 3 T + ER ± DCE, DWI, MRS in EPE (n=17)

Abbreviations: DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; EPE, extraprostatic extension; ER, endorectal coil; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; SS, sample size; T, Tesla.

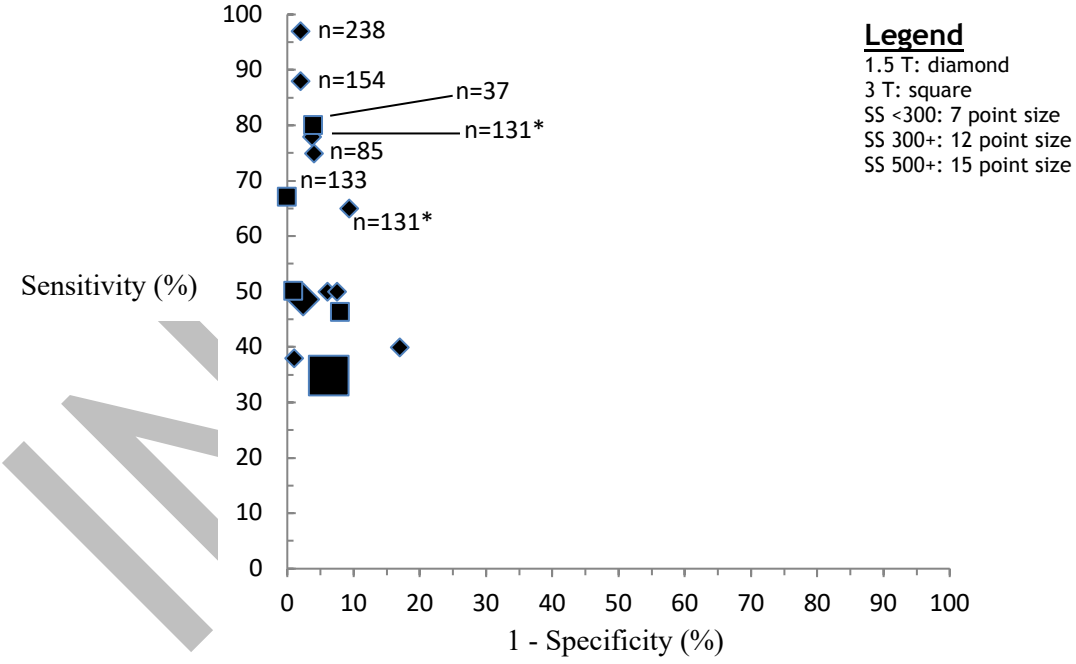


Figure 4-2. MRI Studies of 1.5 T + 3 T + ER ± DCE, DWI, MRS in SVI (n=15)

Abbreviations: DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; ER, endorectal coil; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopic imaging; SS, sample size; SVI, seminal vesicle invasion; T, Tesla.
*Denotes data point from same study population (Soylu et al, 2013).

Table 4-2. Characteristics of Diagnostic Accuracy Studies

Characteristic	No. Data or Studies (%)
EPE	
Field strength	
1.5 T	9 (52.9)
3 T	8 (47.1)
Functional techniques	
MRI	8 (47.1)
MRI + DWI	1 (5.9)
MRI + (DCE and/or MRS)	0 (0)
MRI + DWI + (DCE and/or MRS)	8 (47.1)
Studies of >60% sensitivity	
1.5 T	1 (25.0)
3 T	3 (75.0)
MRI	1 (25.0)
MRI + DWI	0 (0)
MRI + (DCE and/or MRS)	0 (0)
MRI + DWI + (DCE and/or MRS)	3 (75.0)
SVI	
Field strength	
1.5 T	10 (66.7)
3 T	5 (33.3)
Functional techniques	
MRI	9 (60.0)
MRI + DWI	1 (6.7)
MRI + (DCE and/or MRS)	0 (0)
MRI + DWI + (DCE and/or MRS)	5 (33.3)
Studies of >60% sensitivity	
1.5 T	5 (71.4)
3 T	2 (28.6)
MRI	3 (42.9)
MRI + DWI	0 (0)
MRI + (DCE and/or MRS)	0 (0)
MRI + DWI + (DCE and/or MRS)	4 (57.1)

Abbreviations: DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; EPE, extraprostatic extension; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; No., number; SVI, seminal vesicle invasion; T, Tesla.

Table 4-3. Summary of Diagnostic Accuracy

	Sensitivity				Specificity			
Current Guideline	**No. Studies**	**Median**	**Min**	**Max**	**No. Studies**	**Median**	**Min**	**Max**
ER Use								
≥1.5 T for EPE	16	50.0	14.0	90.0	16	91.0	74.0	98.0
≥1.5 T for SVI	13	50.0	34.9	97.0	15	96.0	83.1	100.0
No ER Use								
1.5 T	9	36.2	0	81.3	9	90.3	65.0	97.7
3 T	13	58.3	22.0	92.0	13	86.6	55.2	99.0
De Rooij et al (2015) [52]	**No. Studies**	**Estimate**	**95% CI**		**No. Studies**	**Estimate**	**95% CI**	
EPE	45	0.57	0.49-0.64		45	0.91	0.88-0.93	
SVI	34	0.58	0.47-0.68		34	0.96	0.95-0.97	
Stage T3	38	0.61	0.54-0.67		38	0.88	0.85-0.91	

Abbreviations: CI, confidence interval; EPE, extraprostatic extension; ER, endorectal coil; No., number; SVI, seminal vesicle invasion; T, Tesla.

TNM Stage Classification

There were 21 studies that had information on change in stage classification between routine clinical staging and staging based on MRI ± DCE, DWI, MRS. There were 20 studies that had quantitative information on routine clinical stage at baseline and staging after MRI ± DCE, DWI, MRS [5,6,12,18-21,24,27,28,30,31,33,35,45,53-57] (Appendix 7, Table 6). All studies showed upstaging, except for one study by Tanaka et al (2013) [33]. In this study, a lack of individuals clinically staged as T1 and a notably higher risk study population at patient enrollment may explain the difference in the results between this study and the remaining included studies. Qualitative information was available from a preliminary study prior to the Rud et al (2015) randomized controlled trial that examined 199 patients who had MRI+DWI prior to surgery [41,82]. This population, the basis for the 2015 trial, reported that no patients were downstaged [82].

In a subanalysis, whether MRI ± DCE, DWI, MRS upstaging results were correct was examined using staging at pathology as the comparator (Appendix 8, Table 1). There was a trend toward correct staging by MRI across the seven studies [19-21,35,45,53,56]; however there was a wide range in the proportion of patients correctly staged, between 11% and 85%.

Risk Stratification

There were two studies that provided information on changes to patients' risk stratification category. For both studies, there was a change to a higher risk category after imaging [5,45]. (Appendix 7, Table 7).

Treatment Plan

There were six studies that examined changes to the treatment plan upon MRI ± DCE, DWI, MRS [5,14,41,45,51,53]. All six studies suggested some degree of increased therapy for patients (<1% to 43% of patients). A decrease in the intensity of therapy was shown in two (33.3%) studies [14,51], which occurred for approximately 60% of patients in both studies. Histopathological confirmation was examined in three (50.0%) studies [14,51,53], which showed that imaging-based changes to treatment plans were correct in a minimum of 63% of patients and up to 97% of patients (Appendix 7, Table 8).

Positive Surgical Margins

There were 20 studies that reported on positive surgical margins; one randomized controlled trial [41] and 19 observational studies [10,14,16,19,20,25,26,29,30,33,37,39,51,53,54,56,75,76,79] (Appendix 7, Table 9).

Randomized Controlled Trial

A randomized controlled trial conducted in Norway and published in 2015 was identified that examined whether MRI+DWI prior to robotic-assisted laparoscopic radical prostatectomy would improve surgical outcome determined from positive surgical margins [41]. In this trial, 222 patients were randomized to 1.5 T MRI using a six-channel body matrix plus DWI and 216 patients were randomized to the non-MRI group. The primary study endpoint was the presence of positive surgical margins. The time between biopsy and MRI+DWI was an average of 11 weeks (± 7 weeks). The interpreting radiologist had two years of experience and was not blinded to clinical variables. Reported was preoperative DRE and TRUS clinically staged patients.

The results showed a trend toward fewer positive surgical margins in the MRI+DWI group compared with the non-MRI group; however, the result was not statistically significant (MRI+DWI: 43 [19.4%] versus non-MRI: 49 [22.7%], p=0.4). Although a main study endpoint, the difference between groups was only 4%, which was much less than the anticipated clinically relevant amount of 20% based on sample size estimation parameters. Among patients clinically staged as cT1, there was a statistically significant reduction in positive surgical margins in the MRI+DWI group compared with the non-MRI group, not shown for cT2-3 staged patients (cT1, MRI+DWI: 20 [16.0%] versus non-MRI: 31 [27.2%], p=0.035; cT2-3: MRI+DWI: 23 [23.7%] versus non-MRI: 18 [17.6%], p=0.3) [41] (Appendix 7, Table 9).

Observational Studies

There were 19 observational studies that had information on surgical margin status [10,14,16,19,20,25,26,29,30,33,37,39,51,53,54,56,75,76,79]. There were six observational studies that reported the use of MRI ± DCE, DWI, MRS information to alter the nature of radical prostatectomy surgery [25,29,33,51,53,75]. Among the six observational studies, the range of positive surgical margin status was highly variable (6.7% to 31.9%). In all six observational studies, patients had a low risk of spread defined by PSA levels and the manner in which the imaging results were used to alter surgery was not reported in detail. Variability in the range of values for positive surgical margins was also shown among the 13 observational studies that did not incorporate imaging findings into the surgical practice (4.4% to 43.3%) [10,14,16,19,20,26,30,37,39,54,56,76,79] (Appendix 6, Table 7).

Biochemical Recurrence

There were five studies that examined biochemical recurrence [18,23,54,55,79]. Only two studies examined the independent association imaging to the presence of biochemical recurrence [23,55]. One study showed a statistically significant association between imaging and biochemical recurrence among non-organ-confined disease [23], whereas the other study did not [55] (Appendix 7, Table 10).

Heterogeneity

A priori, a number of variables were considered as potential sources of heterogeneity including field strength, use of an ER, use of functional sequences, PSA level, and sample size. Despite our efforts to account for reasons of heterogeneity, additional sources of heterogeneity existed not accounted for in the current synthesis such as patient clinical differences, methodological variation across studies in the use of functional sequences, the experience level of readers/radiologists, the definition of EPE and SVI at radiology interpretation and its

variability based on MRI ± DCE, DWI, MRS, and the lack of explicit standardization criteria surrounding pathology evaluation (Appendix 8 and 9). We further investigated whether 2 × 2 tables of true positives, false positives, false negatives, and true negatives could be reconstructed. However, due to rounding errors, it was not possible to accurately derive 2 × 2 tables based on published sensitivity, specificity, positive predictive value, and negative predictive value information. Therefore, a meta-analysis was not performed.

DISCUSSION

According to the research questions:

a) What is the performance and diagnostic accuracy of MRI ± DCE, DWI, MRS in men with newly diagnosed biopsy-confirmed prostate cancer who are under consideration for radical treatment?

The performance and diagnostic accuracy of MRI (1.5 T + ER and 3 T ± ER) ± DCE, DWI, MRS in the detection of EPE or SVI was modest in terms of sensitivity (≥50%) and excellent in terms of specificity (>85%), as shown by our current synthesis and systematic review of primary studies. Our synthesis and current systematic review of the literature is consistent with a well-conducted recently published meta-analysis and systematic review that reported overall sensitivity of approximately 50% to 60% and specificity of >85% despite differences in methodology (e.g., years searched), included studies (e.g., use of ER, field strength, use of functional sequences), and analyses.

In our work, above-average sensitivities were achieved in some studies (>60%) suggesting that the current state of MRI ± DCE, DWI, MRS imaging performance may be improved. However, whether elevated sensitivities were achieved due to the addition of functional sequences or particular field strength was not clear owing to too few studies. The current synthesis of the diagnostic accuracy data shows improvements in sensitivity when staging for SVI and improvements in specificity when staging for EPE. Previous work has shown sensitivity and specificity of MRI for local staging of prostate cancer to range from 15% to 100% and 67% to 100%, respectively [62]. Compared with previous work, our synthesis of the contemporary and current published data shows a marked shift toward a higher minimum specificity value in the detection of EPE (74% to 98% versus 49% to 99%; difference: 25%) and a shift toward a higher minimum sensitivity value in the detection of SVI (35% to 97% versus 23% to 80%; difference: 12%) [62]. Higher performance was likely due to analyzing studies according to ER status. Among studies not using an ER, 1.5 T studies had a lower sensitivity than 3 T studies.

Studies of imaging performance by nature lend themselves to designs of single-arm cohort studies, otherwise referred to here as case series. Future studies should consider improvements in study design in terms of patient sampling and study design, standardized use of functional sequences, explicit pathology criteria, and blinding of both radiologists and pathologists. Deficiencies in study quality in the current synthesized data suggest that higher levels of sensitivities may be achieved. Lower sensitivities were shown for studies of 1.5 T MRI ± DCE, DWI, MRS that did not use an ER, whereas imaging performance for studies of 3 T MRI ± DCE, DWI, MRS that did not use an ER was similar to the results obtained for studies that used an ER. This suggests the importance of attention to the technology being used and that a program of quality assurance for centres may be necessary. Additional technical guidance on field strength, use of functional sequences, and ER use is described in Appendix 10. For large studies that showed poor sensitivities (<50%), the reason is not clear. Potential explanations may surround the lack of detail reported for radiology and pathology evaluation and interpretation and the difficulty in distinguishing between focal (a few glands) to very extensive (e.g., 1 cm) EPE [83].

There is some suggestion that intermediate-risk patients had better imaging performance compared with their low-risk counterparts. Based on patients at enrollment, there were no studies that included high-risk patients according to PSA level. The data were examined by PSA level, which can be considered a crude single criterion of tumour risk of spread compared with the routine use of comprehensive tools such as the D'Amico classification and others which use multiple pieces of information including clinical stage (TNM classification), Gleason score, and PSA level [84]. The recommendations are based on the definition of risk used in the current Cancer Care Ontario Prostate Cancer Treatment Pathway [2].

Diagnostic accuracy was examined based on published estimates. From these data, the false-positive fraction (i.e., 1 – specificity) was displayed in scatterplot analysis and the median and range of the specificity was summarized. For EPE, the average false-positive fraction was 9% and for SVI, it was 4% [85]. Our results showed that the minimum time interval between biopsy and MRI was two weeks, although our synthesis did not explore the relationship between time interval and outcomes. Current technical specifications advocate for careful consideration of biological and technical factors that may affect imaging results [1]. A formal evaluation of harms outcomes was out of scope; however, notably, in the included studies, patients with contraindications to MRI were excluded and the current Ontario standard of practice is to screen for suitable patients (e.g., renal function, lack of contraindications).

b) What is the impact of pre-treatment local staging by MRI ± DCE, DWI, MRS on patient outcomes, biochemical recurrence, changes in treatment planning (including nerve-sparing surgery), changes in stage classification, and surgical margin status in men with newly diagnosed biopsy-confirmed prostate cancer who are under consideration for radical treatment?

The impact of pre-treatment local staging by MRI ± DCE, DWI, MRS on clinical and patient outcomes is still not clear. Although studies consistently demonstrated upstaging by MRI ± DCE, DWI, MRS compared with routine clinical staging, this was only correct by pathology in a low number of studies and the magnitude was highly variable (11% to 85%), thus making any conclusions to be drawn imprecise. There was limited information on the outcome of risk stratification category (two studies); however, the results were consistent with the upstaging trend shown for the outcome of TNM staging classification. Changes to treatment plan with MRI ± DCE, DWI, MRS were mixed as they included both an increase and decrease to therapy intensity; however, with a tendency toward more intense therapy across included studies. The correctness of treatment changes was understudied (two studies). Overall, the paucity of data on treatment planning related to MRI ± DCE, DWI, MRS makes it difficult to draw robust conclusions. The proportion of positive surgical margins was examined in two groups stratified by whether the study reported having used the imaging information to alter surgical practice. Examining studies in aggregate, the proportion of positive surgical margins among studies that used MRI ± DCE, DWI, MRS information to alter surgery did not show a benefit as would be demonstrated by lower positive surgical margin values compared with studies that did not use the MRI ± DCE, DWI, MRS information to alter surgery. The studies that reported changes to surgery based on MRI ± DCE, DWI, MRS lacked sufficient details in the manner in which the radiologist's interpretation of tumour stage was translated to surgical practice. The outcome of positive surgical margin status is a complex one, with its clinical impact increasingly scrutinized [86]. Moreover, additional differences between the two stratified groups being compared, including PSA level, pathological Gleason sum, pathological category, surgical modality, year of surgery, and location of surgery are additional factors that need consideration when comparing two groups [87]. The relation between tumour or EPE detected on MRI ± DCE, DWI, MRS to biochemical recurrence was mixed and difficult to decipher due to unclear

reporting and statistical methods. The limitations of the studies included in this systematic review for investigation of clinical outcomes include the small study sizes, the paucity of data, the lack of consistently reported outcomes across studies, and the lack of comparable analysis methods across studies. The complexity involved in the translation of findings on MRI ± DCE, DWI, MRS to changes in surgery, and the complexity in communication between radiologists and urologists likely contributes to the difficulty of investigation [41].

The tool of GRADE (Grading of Recommendations Assessment, Development and Evaluation) was not used as most studies identified were observational in nature. According to GRADE, an observational study would signal a 'low' quality assessment score. One randomized controlled trial was identified, and it served to inform the recommendations, particularly with respect to surgical margins.

CONCLUSIONS

In summary, the literature search revealed predominately diagnostic accuracy studies on MRI ± DCE, DWI, MRS. Modest imaging performance using MRI (1.5 T + ER and 3 T ± ER) ± DCE, DWI, MRS in the detection of EPE and SVI for patients scheduled to undergo radical prostatectomy was demonstrated (sensitivities: 50% to 60%, specificities: >85%). Additional technical guidance when using MRI is described in Appendix 10. Scarcity of data for clinical and patient outcomes limited synthesis and examination. Future research is needed in the form of high-quality diagnostic accuracy studies and health outcomes studies using randomized controlled trial designs to help inform evidence-based recommendations and practice. Our recommendations are tempered due to the limitations of included studies and our recommendations are best used in the context of the current Cancer Care Ontario Prostate Cancer Diagnosis Pathway [2] and the current minimum requirements given by Prostate Imaging and Reporting and Data Systems Version 2 (PI-RADS v2) [1] (Appendix 10).

Magnetic Resonance Imaging for Pre-Treatment Local Staging of Prostate Cancer

Section 5: Internal and External Review

INTERNAL REVIEW

The guideline was evaluated by the Guideline Development Group (GDG) Expert Panel and the Program in Evidence-Based Care (PEBC) Report Approval Panel (RAP) (Appendix 2). The results of these evaluations and the Working Group's responses are described below.

Expert Panel Review and Approval

Of the 10 members of the GDG Expert Panel, nine members cast votes, for a total of 90% response between October and November 2015. Of the nine members that cast votes, eight approved the document (88.9%). The main comments from the Expert Panel and the Working Group's responses are summarized in Table 5-1.

Table 5-1. Summary of the Working Group's responses to comments from the Expert Panel.

Comments	Responses
1. Distinguish between focal to very extensive extraprostatic extension in standard pathology reporting and in magnetic resonance imaging (MRI) detection.	Added suggested wording and references to Section 4, under Discussion.
2. Change recommendations to: "should not yet be considered the standard of care" and "may be considered in high-risk patients".	Incorporated high-risk patients into Recommendation 1 as part of consensus expert clinical opinion.
3. Add bulleted point to recommendations about the delivery of non-surgical treatments.	Considered to be out-of-scope.

RAP Review and Approval

Three RAP members, including the PEBC Director, reviewed this document in October 2015. The results of RAP were: one approval (October 18, 2015), and two conditional approvals (October 23 and 29, 2015). The main comments from the RAP and the Working Group's responses are summarized in Table 5-2.

Table 5-2. Summary of the Working Group's responses to comments from RAP.

Comments	Responses
1. Better define low-, intermediate- and high-risk patients.	Provided definition of risk groups based on Cancer Care Ontario's Prostate Cancer Treatment Pathway in Qualifying Statements for Recommendation 1 and in Section 4, under Discussion.
2. Recommendations: change "may be considered" to more explicit wording such as "is a reasonable option" or "is an option", or "should be considered".	Changed Recommendation 1 to "is a reasonable option".
3. Give a sense of where the technology fits in?	Added paragraph to Section 3, under Background for Guideline.
4. The needed level of experience of the interpreting radiologist is debatable.	Deleted wording.
5. Formulating recommendations: any renal risks? Other harms from upstaging?	Added wording to Section 4, under Discussion.

6. Strengthen the link between the evidence and recommendations.	Added wording in Section 4, under Discussion, for the need for more high-quality evidence.
7. Can authors address what is meant by "modest" sensitivity?	Added reference to Section 4, under Methods.

EXTERNAL REVIEW

External Review by Ontario Clinicians and Other Experts

Targeted Peer Review

Six targeted peer reviewers from Ontario and the United States who are considered to be clinical and/or methodological experts on the topic were identified by the Working Group. Four individuals agreed to be the reviewers (Appendix 3), with four responses received. Results of the feedback survey are summarized in Table 5-3. The comments from targeted peer reviewers and the Working Group's responses are summarized in Table 5-4.

Table 5-3. Responses to nine items on the targeted peer reviewer questionnaire.

	Reviewer Ratings (N=<x>)				
Question	Lowest Quality (1)	(2)	(3)	(4)	Highest Quality (5)
1. Rate the guideline development methods.				1	3
2. Rate the guideline presentation.				1	3
3. Rate the guideline recommendations.				2	2
4. Rate the completeness of reporting.				2	2
5. Does this document provide sufficient information to inform your decisions? If not, what areas are missing?				2	2
6. Rate the overall quality of the guideline report.				1	3
	Strongly Disagree (1)	(2)	Neutral (3)	(4)	Strongly Agree (5)
7. I would make use of this guideline in my professional decisions.				2	2
8. I would recommend this guideline for use in practice.				2	2
9. What are the barriers or enablers to the implementation of this guideline report?	-Magnetic resonance imaging availability -Quality assurance -Clinician uptake				

Table 5-4. Responses to comments from targeted peer reviewers.

Comments	Responses
1. Preferable to include cost considerations?	Out-of-scope.
2. How can Prostate Imaging and Reporting and Data Systems (PI-RADS) be incorporated?	Guideline speaks to multiparametric magnetic resonance imaging (mpMRI); therefore, PI-RADS Version 2 has been incorporated. See Appendix 10 for mpMRI and technical specifications.
3. Are you to address MRI in diagnosis and surveillance?	Out-of-scope, see PEBC reports #27-2 and #17-9. https://www.cancercare.on.ca/cms/One.aspx?portalId=1377&pageId=10144

4. Can you suggest an optimum time post-biopsy to perform MRI?	We make no specific recommendation on this. See Discussion and Appendix 10 for mpMRI and technical specifications.

Professional Consultation

Feedback was obtained through a brief online survey of healthcare professionals and other stakeholders who are the intended users of the guideline. All individuals in the PEBC database categorized as urologists, radiologists, radiation oncologists, and general practitioners were contacted by email to inform them of the survey. There were 209 individuals contacted: 181 in Ontario versus 28 outside Ontario. There were 33 (15.8%) responses received. Of the non-participants, 23 (11.0%) stated that they did not have interest in this area or were unavailable to review this guideline at the time. The results of the feedback survey are summarized in Table 5-5. The main comments from the consultation and the Working Group's responses are summarized in Table 5-6.

Table 5-5. Responses to four items on the professional consultation survey.

	Number (%)				
General Questions: Overall Guideline Assessment	Lowest Quality (1)	(2)	(3)	(4)	Highest Quality (5)
1. Rate the overall quality of the guideline report.			2 (6.1)	18 (54.5)	13 (39.4)
	Strongly Disagree (1)	(2)	(3)	(4)	Strongly Agree (5)
2. I would make use of this guideline in my professional decisions.		2 (6.1)	4 (12.1)	14 (42.4)	13 (39.4)
3. I would recommend this guideline for use in practice.			6 (18.2)	13 (39.4)	14 (42.4)
4. What are the barriers or enablers to the implementation of this guideline report?	-Magnetic resonance imaging access & timely availability -Quality assurance -Clinician/health administrators uptake -Clinical training/expertise				

Table 5-6. Modifications/Actions taken/Responses regarding main written comments from professional consultants.

Comments	**Responses**
1. Are there technical specifications to be included? Also, additional guidance around 1.5 T use?	Minimum requirements are outlined in Prostate Imaging and Reporting and Data Systems Version 2 (PI-RADS v2), and have been included in Appendix 10, multiparametric magnetic resonance imaging and technical specifications.
2. What would a quality assurance program look like, and what would its implementation look like?	Implementation considerations have been outlined, see page 5. Additional implementation details are out-of-scope for PEBC.
3. Radiation treatment planning for focal dose escalation should at least be mentioned.	Out-of-scope.
4. Can there be more detail around inter-observer agreement, or diagnostic accuracy of false negatives, false	Yes, please see Discussion. Otherwise, inter-observer agreement was not part of the scope.

positives, and additional clarification around “time to MRI after biopsy”.	
5. More clear and direct qualifying statement beginning at “Caution…”	Re-phrased this sentence.
6. Please mention Appendix 10 in the discussion. This is important information not to be overlooked.	Yes, it has been stated throughout the document.
7. Typo in Appendix 2, should read >20.	Yes, corrected.

Publication Peer Review Feedback

Policy Review

CONCLUSION

The final guideline recommendations contained in Section 2 and summarized in Section 1 reflect the integration of feedback obtained through the external review processes with the document as drafted by the GDG Working Group and approved by the GDG Expert Panel and the PEBC RAP.

References

1. American College of Radiology. Prostate Imaging and Reporting and Data System 2015 Version 2 [cited March 11, 2016]. Available from: http://www.acr.org/~/media/ACR/Documents/PDF/QualitySafety/Resources/PIRADS/PIRADS%20V2.pdf.

2. Finelli A, Loblaw A, Toi A, Grabowski A, Shenfield C, Booth C, et al. Prostate Cancer Treatment Pathway Version 2013.01 [cited March 11, 2016]. Available from: https://https://www.cancercareontario.ca/en/pathway-maps/prostate-cancer.

3. Schieda N, Quon JS, Lim C, El-Khodary M, Shabana W, Singh V, et al. Evaluation of the European Society of Urogenital Radiology (ESUR) PI-RADS scoring system for assessment of extra-prostatic extension in prostatic carcinoma. Eur J Radiol. 2015;84(10):1843-8.

4. Ghafoori M, Alavi M, Shakiba M, Hoseini K. The value of prostate MRI with endorectal coil in detecting seminal vesicle involvement in patients with prostate cancer. Iran J Radiol. 2015;12(1):e14556.

5. Counago F, Recio M, del Cerro E, Cerezo L, Diaz Gavela A, Marcos FJ, et al. Role of 3.0 T multiparametric MRI in local staging in prostate cancer and clinical implications for radiation oncology. Clin Transl Oncol. 2014;16(11):993-9.

6. Gupta RT, Faridi KF, Singh AA, Passoni NM, Garcia-Reyes K, Madden JF, et al. Comparing 3-T multiparametric MRI and the Partin tables to predict organ-confined prostate cancer after radical prostatectomy. Urol Oncol. 2014;32(8):1292-9.

7. Kan RW, Kan CF, Ho LY, Chan SW. Pre-operative tumor localization and evaluation of extra-capsular extension of prostate cancer: how misleading can it be? Urol J. 2014;11(3):1615-9.

8. Kim CK, Park SY, Park JJ, Park BK. Diffusion-weighted MRI as a predictor of extracapsular extension in prostate cancer. ACR Am J Roentgenol. 2014;202(3):W270-W6.

9. Kim JY, Kim SH, Kim YH, Lee HJ, Kim MJ, Choi MS. Low-risk prostate cancer: the accuracy of multiparametric MR imaging for detection. Radiology. 2014;271(2):435-44.

10. Kitamura K, Muto S, Yokota I, Hoshimoto K, Kaminaga T, Noguchi T, et al. Feasibility of multiparametric prostate magnetic resonance imaging in the detection of cancer distribution: histopathological correlation with prostatectomy specimens. Prostate Int. 2014;2(4):188-95.

11. Lawrence EM, Gallagher FA, Barrett T, Warren AY, Priest AN, Goldman D, et al. Preoperative 3-T diffusion-weighted MRI for the qualitative and quantitative assessment of extracapsular extension in patients with intermediate- or high-risk prostate cancer. ACR Am J Roentgenol. 2014;203(3):W280-6.

12. Lista F, Gimbernat H, Caceres F, Rodriguez-Barbero JM, Castillo E, Angulo JC. Multiparametric magnetic resonance imaging for the assessment of extracapsular invasion and other staging parameters in patients with prostate cancer candidates for radical prostatectomy. Actas Urol Esp. 2014;38(5):290-7.

13. Otto J, Thormer G, Seiwerts M, Fuchs J, Garnov N, Franz T, et al. Value of endorectal magnetic resonance imaging at 3T for the local staging of prostate cancer. Rofo. 2014;186(8):795-802.

14. Park BH, Jeon HG, Jeong BC, Seo SI, Lee HM, Choi HY, et al. Influence of magnetic resonance imaging in the decision to preserve or resect neurovascular bundles at robotic assisted laparoscopic radical prostatectomy. J Urol. 2014;192(1):82-8.

15. Roethke M, Kaufmann S, Kniess M, Ketelsen D, Claussen CD, Schlemmer HP, et al. Seminal vesicle invasion: accuracy and analysis of infiltration patterns with high-spatial resolution T2-weighted sequences on endorectal magnetic resonance imaging. Urol Int. 2014;92(3):294-9.

16. Song SH, Pak S, Park S, Song C, Jeong IG, Choi HJ, et al. Predictors of unfavorable disease after radical prostatectomy in patients at low risk by D'Amico criteria: role of multiparametric magnetic resonance imaging. J Urol. 2014;192(2):402-8.

17. Styles C, Ferris N, Mitchell C, Murphy D, Frydenberg M, Mills J, et al. Multiparametric 3T MRI in the evaluation of intraglandular prostate cancer: correlation with histopathology. J Med Imaging Radiat Oncol. 2014;58(4):439-48.

18. Armitage JN, Igali L, Ball RY, Mills RD, Kumar V. Comparison of body-array MRI and Partin tables for predicting extracapsular prostate cancer. J Clin Urol. 2013;6(1):50-4.

19. Cerantola Y, Valerio M, Kawkabani Marchini A, Meuwly JY, Jichlinski P. Can 3T multiparametric magnetic resonance imaging accurately detect prostate cancer extracapsular extension? Can Urol Assoc J. 2013;7(11-12):E699-703.

20. Hegde JV, Chen MH, Mulkern RV, Fennessy FM, D'Amico AV, Tempany CM. Preoperative 3-Tesla multiparametric endorectal magnetic resonance imaging findings and the odds of upgrading and upstaging at radical prostatectomy in men with clinically localized prostate cancer. Int J Radiat Oncol Biol Phys. 2013;85(2):e101-7.

21. Hole KH, Axcrona K, Lie AK, Vlatkovic L, Geier OM, Brennhovd B, et al. Routine pelvic MRI using phased-array coil for detection of extraprostatic tumour extension: accuracy and clinical significance. Eur Radiol. 2013;23(4):1158-66.

22. Isebaert S, Van Den Bergh L, Haustermans K, Joniau S, Lerut E, De Wever L, et al. Multiparametric MRI for prostate cancer localization in correlation to whole-mount histopathology. J Magn Reson Imaging. 2013;37(6):1392-401.

23. Jeong IG, Lim JH, You D, Kim MH, Choi HJ, Kim JK, et al. Incremental value of magnetic resonance imaging for clinically high risk prostate cancer in 922 radical prostatectomies. J Urol. 2013;190(6):2054-60.

24. Johnston R, Wong LM, Warren A, Shah N, Neal D. The role of 1.5 Tesla magnetic resonance imaging in staging prostate cancer. ANZ J Surg. 2013;83(4):234-8.

25. Nepple KG, Rosevear HM, Stolpen AH, Brown JA, Williams RD. Concordance of preoperative prostate endorectal MRI with subsequent prostatectomy specimen in high-risk prostate cancer patients. Urol Oncol. 2013;31(5):601-6.

26. Pak S, Park S, Ryu J, Hong S, Song SH, You D, et al. Preoperative factors predictive of posterolateral extracapsular extension after radical prostatectomy. Korean J Urol. 2013;54(12):824-9.

27. Porcaro AB, Borsato A, Romano M, Sava T, Ghimenton C, Migliorini F, et al. Accuracy of preoperative endo-rectal coil magnetic resonance imaging in detecting clinical under-staging of localized prostate cancer. World J Urol. 2013;31(5):1245-51.

28. Pugh TJ, Frank SJ, Achim M, Kuban DA, Lee AK, Hoffman KE, et al. Endorectal magnetic resonance imaging for predicting pathologic T3 disease in Gleason score 7 prostate cancer: Implications for prostate brachytherapy. Brachytherapy. 2013;12(3):204-9.

29. Roethke MC, Lichy MP, Kniess M, Werner MK, Claussen CD, Stenzl A, et al. Accuracy of preoperative endorectal MRI in predicting extracapsular extension and influence on neurovascular bundle sparing in radical prostatectomy. World J Urol. 2013;31(5):1111-6.

30. Renard-Penna R, Roupret M, Comperat E, Ayed A, Coudert M, Mozer P, et al. Accuracy of high resolution (1.5 tesla) pelvic phased array magnetic resonance imaging (MRI) in staging prostate cancer in candidates for radical prostatectomy: results from a prospective study. Urol Oncol. 2013;31(4):448-54.

31. Somford DM, Hamoen EH, Futterer JJ, Van Basten JP, Hulsbergen-Van De Kaa CA, Vreuls W, et al. The predictive value of endorectal 3 tesla multiparametric magnetic resonance imaging for extraprostatic extension in patients with low, intermediate and high risk prostate cancer. J Urol. 2013;190(5):1728-34.

32. Soylu FN, Peng Y, Jiang Y, Wang S, Schmid-Tannwald C, Sethi I, et al. Seminal vesicle invasion in prostate cancer: evaluation by using multiparametric endorectal MR imaging. Radiology. 2013;267(3):797-806.

33. Tanaka K, Shigemura K, Muramaki M, Takahashi S, Miyake H, Fujisawa M. Efficacy of using three-tesla magnetic resonance imaging diagnosis of capsule invasion for decision-making about neurovascular bundle preservation in robotic-assisted radical prostatectomy. Korean J Urol. 2013;54(7):437-41.

34. Turkbey B, Mani H, Aras O, Ho J, Hoang A, Rastinehad AR, et al. Prostate cancer: can multiparametric MR imaging help identify patients who are candidates for active surveillance? Radiology. 2013;268(1):144-52.

35. Tsao CW, Lin MH, Wu ST, Meng E, Tang SH, Chen HI, et al. Combining prostrate-specific antigen and Gleason score increases the diagnostic power of endorectal coil magnetic resonance imaging in prostate cancer pathological stage. J Chin Med Assoc. 2013;76(1):20-4.

36. Kim BS, Kim TH, Kwon TG, Yoo ES. Comparison of pelvic phased-array versus endorectal coil magnetic resonance imaging at 3 Tesla for local staging of prostate cancer. Yonsei Med J. 2012;53(3):550-6.

37. Novis MI, Baroni RH, Cerri LM, Mattedi RL, Buchpiguel CA. Clinically low-risk prostate cancer: evaluation with transrectal doppler ultrasound and functional magnetic resonance imaging. Clinics (Sao Paulo). 2011;66(1):27-34.

38. Turkbey B, Mani H, Shah V, Rastinehad AR, Bernardo M, Pohida T, et al. Multiparametric 3T prostate magnetic resonance imaging to detect cancer: histopathological correlation using prostatectomy specimens processed in customized magnetic resonance imaging based molds. J Urol. 2011;186(5):1818-24.

39. Lee SH, Park KK, Choi KH, Lim BJ, Kim JH, Lee SW, et al. Is endorectal coil necessary for the staging of clinically localized prostate cancer? Comparison of non-endorectal versus endorectal MR imaging. World J Urol. 2010;28(6):667-72.

40. Turkbey B, Pinto PA, Mani H, Bernardo M, Pang Y, McKinney YL, et al. Prostate cancer: value of multiparametric MR imaging at 3 T for detection--histopathologic correlation. Radiology. 2010;255(1):89-99.

41. Rud E, Baco E, Klotz D, Rennesund K, Svindland A, Berge V, et al. Does preoperative magnetic resonance imaging reduce the rate of positive surgical margins at radical prostatectomy in a randomised clinical trial? Eur Urol. 2015;68(3):487-96.

42. Orczyk C, Rosenkrantz AB, Deng FM, Melamed J, Babb J, Wysock J, et al. A prospective comparative analysis of the accuracy of HistoScanning and multiparametric magnetic resonance imaging in the localization of prostate cancer among men undergoing radical prostatectomy. Urol Oncol. 2016;34(1):3.e1-8.

43. Kayat Bittencourt L, Litjens G, Hulsbergen-van de Kaa CA, Turkbey B, Gasparetto EL, Barentsz JO. Prostate cancer: the European Society of Urogenital Radiology Prostate Imaging Reporting and Data System criteria for predicting extraprostatic extension by using 3-T multiparametric MR imaging. Radiology. 2015;276(2):479-89.

44. Boesen L, Chabanova E, Logager V, Balslev I, Mikines K, Thomsen HS. Prostate cancer staging with extracapsular extension risk scoring using multiparametric MRI: a correlation with histopathology. Eur Radiol. 2015;25(6):1776-85.

45. Counago F, Del Cerro E, Diaz-Gavela AA, Marcos FJ, Recio M, Sanz-Rosa D, et al. Tumor staging using 3.0 T multiparametric MRI in prostate cancer: impact on treatment decisions for radical radiotherapy. Springerplus. 2015;4:789.

46. de Cobelli O, Terracciano D, Tagliabue E, Raimondi S, Bottero D, Cioffi A, et al. Predicting pathological features at radical prostatectomy in patients with prostate cancer eligible for active surveillance by multiparametric magnetic resonance imaging. PLoS ONE [Electronic Resource]. 2015;10(10):e0139696.

47. Feng TS, Sharif-Afshar AR, Wu J, Li Q, Luthringer D, Saouaf R, et al. Multiparametric MRI improves accuracy of clinical nomograms for predicting extracapsular extension of prostate cancer. Urology. 2015;86(2):332-7.

48. Razi A, Parizi MK, Kazemeini SM, Abedi A. A prospective study of the efficacy of magnetic resonance spectroscopy imaging for predicting locally advanced prostate cancer. Turk J Urol. 2015;41(2):67-72.

49. Reisaeter LA, Futterer JJ, Halvorsen OJ, Nygard Y, Biermann M, Andersen E, et al. 1.5-T multiparametric MRI using PI-RADS: a region by region analysis to localize the index-tumor of prostate cancer in patients undergoing prostatectomy. Acta Radiol. 2015;56(4):500-11.

50. Woo S, Cho JY, Kim SY, Kim SH. Extracapsular extension in prostate cancer: added value of diffusion-weighted MRI in patients with equivocal findings on T2-weighted imaging. AJR Am J Roentgenol. 2015;204(2):W168-75.

51. McClure TD, Margolis DJ, Reiter RE, Sayre JW, Thomas MA, Nagarajan R, et al. Use of MR imaging to determine preservation of the neurovascular bundles at robotic-assisted laparoscopic prostatectomy. Radiology. 2012;262(3):874-83.

52. de Rooij M, Hamoen EH, Witjes JA, Barentsz JO, Rovers MM. Accuracy of magnetic resonance imaging for local staging of prostate cancer: a diagnostic meta-analysis. Eur Urol 2015.http://dx.doi.org/10.1016/j.eururo.2015.07.029 (in press).

53. Panebianco V, Salciccia S, Cattarino S, Minisola F, Gentilucci A, Alfarone A, et al. Use of multiparametric MR with neurovascular bundle evaluation to optimize the oncological and functional management of patients considered for nerve-sparing radical prostatectomy. J Sex Med. 2012;9(8):2157-66.

54. Ploussard G, Xylinas E, Durand X, Ouzaid I, Allory Y, Bouanane M, et al. Magnetic resonance imaging does not improve the prediction of misclassification of prostate cancer patients eligible for active surveillance when the most stringent selection criteria are based on the saturation biopsy scheme. BJU Int. 2011;108(4):513-7.

55. Park JJ, Kim CK, Park SY, Park BK, Lee HM, Cho SW. Prostate cancer: Role of pretreatment multiparametric 3-T MRI in predicting biochemical recurrence after radical prostatectomy. ACR Am J Roentgenol. 2014;202(5):W459-W65.

56. Brown JA, Rodin DM, Harisinghani M, Dahl DM. Impact of preoperative endorectal MRI stage classification on neurovascular bundle sparing aggressiveness and the radical prostatectomy positive margin rate. Urol Oncol. 2009;27(2):174-9.

57. Cirillo S, Petracchini M, Bona CM, Durando S, Ortega C, Vormola R, et al. Comparison of endorectal magnetic resonance imaging, clinical prognostic factors and nomograms in the local staging of prostate cancer patients treated with radiotherapy. Tumori. 2008;94(1):65-9.

58. Browman GP, Levine MN, Mohide EA, Hayward RS, Pritchard KI, Gafni A, et al. The practice guidelines development cycle: a conceptual tool for practice guidelines development and implementation. J Clin Oncol. 1995;13(2):502-12.

59. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.

60. Canadian Cancer Society's Advisory. Committee on Cancer Statistics. Canadian Cancer Statistics 2015. Toronto OCCS.

61. Yacoub JH, Oto A, Miller FH. MR Imaging of the prostate. Radiol Clin North Am. 2014;52(4):811-37.

62. Zukotynski K, Haider MA. Imaging in prostate carcinoma. Hematol Oncol Clin North Am. 2013;27(6):1163-87.

63. Thompson J, Lawrentschuk N, Frydenberg M, Thompson L, Stricker P. The role of magnetic resonance imaging in the diagnosis and management of prostate cancer. BJU Int. 2013;112(SUPPL. 2):6-20.

64. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7:10.

65. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.

66. Sterne JAC HJ, Reeves BC on behalf of the development group for ACROBATNRSI. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBATNRSI), Version 1.0.0, 24 September 2014. Available from http://www.riskofbias.info.

67. Szklo M, Nieto JF. Chapter 8 Quality Assurance and Control. Epidemiology Beyond the Basics. Gaithersburg, Maryland, USA: Aspen Publishers, Inc.; 2000. p. 343-404.

68. Silva RC, Sasse AD, Matheus WE, Ferreira U. Magnetic resonance image in the diagnosis and evaluation of extra-prostatic extension and involvement of seminal vesicles of prostate cancer: a systematic review of literature and meta-analysis. Int Braz J Urol. 2013;39(2):155-66.

69. Haghighi M, Shah S, Taneja SS, Rosenkrantz AB. Prostate cancer: diffusion-weighted imaging versus dynamic-contrast enhanced imaging for tumor localization - a meta-analysis. J Comput Assist Tomogr. 2013;37(6):980-8.

70. Wu LM, Xu JR, Ye YQ, Lu Q, Hu JN. The clinical value of diffusion-weighted imaging in combination with T2-weighted imaging in diagnosing prostate carcinoma: a systematic review and meta-analysis. AJR Am J Roentgenol. 2012;199(1):103-10.

71. Umbehr M, Bachmann LM, Held U, Kessler TM, Sulser T, Weishaupt D, et al. Combined magnetic resonance imaging and magnetic resonance spectroscopy imaging in the diagnosis of prostate cancer: a systematic review and meta-analysis. Eur Urol. 2009;55(3):575-90.

72. Hovels AM, Heesakkers RA, Adang EM, Jager GJ, Strum S, Hoogeveen YL, et al. The diagnostic accuracy of CT and MRI in the staging of pelvic lymph nodes in patients with prostate cancer: a meta-analysis. Clin Radiol. 2008;63(4):387-95.

73. Mowatt G, Scotland G, Boachie C, Cruickshank M, Ford JA, Fraser C, et al. The diagnostic accuracy and cost-effectiveness of magnetic resonance spectroscopy and enhanced magnetic resonance imaging techniques in aiding the localisation of prostate abnormalities for biopsy: a systematic review and economic evaluation. Health Technol Assess. 2013;17(20).

74. National Institute for Health and Care Excellence. Prostate cancer: diagnosis and treatment 2014 [updated January 2014; cited Accessed March 11, 2016]. Available from: https://www.nice.org.uk/guidance/cg175.

75. Yao A, Iwamoto H, Masago T, Morizane S, Honda M, Sejima T, et al. The role of staging MRI in predicting apical margin positivity for robot-assisted laparoscopic radical prostatectomy. Urol Int. 2014;93(2):182-8.

76. Guzzo TJ, Resnick MJ, Canter DJ, Bivalacqua TJ, Rosen MA, Bergey MR, et al. Endorectal T2-weighted MRI does not differentiate between favorable and adverse pathologic features in men with prostate cancer who would qualify for active surveillance. Urol Oncol. 2012;30(3):301-5.

77. Jung AJ, Coakley FV, Shinohara K, Carroll PR, Kurhanewicz J, Cowan JE, et al. Local staging of prostate cancer: comparative accuracy of T2-weighted endorectal MR imaging and transrectal ultrasound. Clin Imaging. 2012;36(5):547-52.

78. Hwii Ko Y, Jae Sung D, Gu Kang S, Ho Kang S, Gu Lee J, Jong Kim J, et al. The predictability of T3 disease in staging MRI following prostate biopsy decreases in patients with high initial PSA and Gleason score. Asian J Androl. 2011;13(3):487-93.

79. Lee HW, Seo SI, Jeon SS, Lee HM, Choi HY. Can we predict real T3 stage prostate cancer in patients with clinical T3 (cT3) disease before radical prostatectomy? Yonsei Med J. 2010;51(5):700-7.

80. Zhang J, Hricak H, Shukla-Dave A, Akin O, Ishill NM, Carlino LJ, et al. Clinical stage T1c prostate cancer: evaluation with endorectal MR imaging and MR spectroscopic imaging. Radiology. 2009;253(2):425-34.

81. Junker D, Quentin M, Nagele U, Edlinger M, Richenberg J, Schaefer G, et al. Evaluation of the PI-RADS scoring system for mpMRI of the prostate: a whole-mount step-section analysis. World J Urol. 2015;33(7):1023-30.

82. Rud E, Klotz D, Rennesund K, Baco E, Johansen TE, Diep LM, et al. Preoperative magnetic resonance imaging for detecting uni- and bilateral extraprostatic disease in patients with prostate cancer. World J Urol. 2015;33(7):1015-21.

83. Magi-Galluzzi C, Evans AJ, Delahunt B, Epstein JI, Griffiths DF, van der Kwast TH, et al. International Society of Urological Pathology (ISUP) Consensus Conference on Handling and Staging of Radical Prostatectomy Specimens. Working group 3: extraprostatic extension, lymphovascular invasion and locally advanced disease. Mod Pathol. 2011;24(1):26-38.

84. National Comprehensive Cancer Network Version 1 2015. Prostate Cancer [cited December 9, 2015]. Available from: http://www.nccn.org/patients/guidelines/prostate/files/assets/common/downloads/files/prostate.pdf.

85. Macaskill P GC, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane, Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration Afhsco.

86. Yossepowitch O, Briganti A, Eastham JA, Epstein J, Graefen M, Montironi R, et al. Positive surgical margins after radical prostatectomy: a systematic review and contemporary update. Eur Urol. 2014;65(2):303-13.

87. Preston MA, Breau RH, Lantz AG, Morash C, Gerridzen RG, Doucette S, et al. The association between nerve sparing and a positive surgical margin during radical prostatectomy. Urol Oncol. 2015;33(1):18 e1-6.

88. Feng TS, Sharif-Afshar AR, Smith SC, Miller J, Nguyen C, Li Q, et al. Multiparametric magnetic resonance imaging localizes established extracapsular extension of prostate cancer. Urol Oncol. 2015;33(3):109.e15-22.

Appendix 1. Glossary of Terms

Functional sequences
Includes dynamic contrast-enhanced imaging (DCE), diffusion-weighted imaging (DWI), and proton magnetic resonance spectroscopic imaging (MRS). One or more of these functional sequences are used with T2-weighted magnetic resonance imaging.

Magnetic resonance imaging
Refers to T2-weighted magnetic resonance imaging.

Multiparametric magnetic resonance imaging
Is a form of MRI that typically includes two or more functional sequences; however, multiparametric magnetic resonance imaging is also often described when only one additional functional sequence is considered in addition to T2-weighted magnetic resonance imaging.

PI-RADS v2
Refers to Prostate Imaging and Reporting and Data Systems Version 2. This document specifies the minimum requirements for MRI and functional sequences, scanner and ER use, and time to imaging for staging post-biopsy.

Appendix 2. Definitions of Low, Intermediate, and High Risk

Low Risk (all of)
Prostate specific antigen <10 ng/mL
Gleason score ≤6
Clinical stage T1-T2a
Asymptomatic for metastasis

Intermediate Risk (i.e., neither low risk nor high risk)
Prostate specific antigen 10 to 20 ng/mL
Gleason score 7
Clinical stage T2b
Asymptomatic for metastasis

High Risk (any of)
Prostate specific antigen >20 ng/mL
Gleason score ≥6
Clinical stage T2c-3a
Symptomatic for metastasis

***
Reference: Finelli et al, 2013 [2]

Appendix 3. Membership

Members of the Magnetic Resonance Imaging in Pre-Treatment Staging of Prostate Cancer Guideline Development Group (Working Group, Expert Panel), Report Approval Panel, Target Reviewers and their Conflict of Interest Declaration.

Table 1. Members of the Guideline Development Group

Name	Affiliation	Declarations of interest
Working Group		
Masoom Haider Working Group Chair (Radiologist)	Sunnybrook Health Sciences Centre, Toronto, ON	Presentations at numerous scientific meetings about magnetic resonance imaging (MRI), advocates for its use in planning and staging in prostate cancer, financial incentives
Antonio Finelli (Urologist)	Princess Margaret Hospital, Toronto, ON	None declared
Chris Morash (Urologist)	The Ottawa Hospital, Ottawa, ON	None declared
Scott Morgan (Radiation Oncologist)	Ottawa Hospital Research Institute, Ottawa, ON	None declared
Nick Power (Urologist)	Centre for Translational Cancer Research, London, ON	None declared
Nicola Schieda (Radiologist)	Ottawa Hospital Research Institute, Ottawa, ON	None declared
Jennifer Salerno (Health Research Methodologist)	McMaster University Hamilton, ON	None declared
Expert Panel		
Neil Fleshner (Urologist)	Princess Margaret Hospital Toronto, ON	None declared
Glenn Bauman (Radiation Oncologist)	London Health Sciences Centre London, ON	Receives funding and trial co-investigator for multiparametric MRI and prostate cancer; published on MRI and radiotherapy planning
Joseph Chin (Urologist)	London Health Sciences Centre London, ON	None declared
Bobby Shayegan (Urologist)	St. Joseph's Hospital Hamilton, ON	None declared
Rodney Breau (Urologist)	The Ottawa Hospital Ottawa, ON	None declared
Julian Dobranowski (Radiologist)	St. Joseph's Hospital Hamilton, ON	None declared
Kartik Jhaveri (Radiologist)	University Health Network, Mount Sinai Hospital, Women's College Hospital, University of Toronto Toronto, ON	None declared

Sangeet Ghai (Radiologist)	University Health Network Toronto, ON	Support received from imaging company and grant application activity in prostate imaging
Michael Brundage (Radiation Oncologist)	Kingston General Hospital Kingston, ON	None declared
Theo van der Kwast (Pathologist)	Toronto General Hospital Toronto, ON	None declared
Report Approval Panel		
Melissa Brouwers (PEBC Director)	McMaster University Hamilton, ON	None declared
Craig Earle (Methodology Expertise)	Institute for Clinical Evaluative Sciences Toronto, ON	None declared
Shailendra Verma (Methodology Expertise)	The Ottawa Hospital Regional Cancer Centre Ottawa, ON	None declared
Targeted Peer Review		
Laurence Klotz	Sunnybrook Health Sciences Centre Toronto, ON	Involved in relevant business entities, additional relevant support, responsibilities and trial PI involvement
Dan Margolis	University of California, Los Angeles Los Angeles, CA, USA	Support from Siemens Medical Systems (institutional grant) and trial PI (Prospective Assessment of Image Registration for the Diagnosis of Prostate Cancer)
Aytek Oto	The University of Chicago Medicine Chicago, IL, USA	Consulting support from Profound Healthcare and support from Philips Healthcare
Joshua Tjong	Sault Ste Marie Hospital Sault Ste Marie, ON	None declared

Conflict of Interest

In accordance with the Program in Evidence-Based Care (PEBC) Conflict of Interest (COI) Policy, the guideline authors, Expert Panel members, and internal and external reviewers were asked to disclose potential conflicts of interest. The COI declared above did not disqualify any individuals from performing their designated role in the development of this guideline, in accordance with the PEBC COI Policy. To obtain a copy of the policy, please contact the PEBC office by email at ccopgi.mcmaster.ca

Appendix 4. AMSTAR Results

Table 1. Critical Appraisal Using AMSTAR

Criteria of Fulfillment Using AMSTAR	NICE (2014) [74]	De Rooij (2015) [52]
Q1: Research question and inclusion criteria stated?	Yes (1)	Yes (1)
Q2: Duplicate verification?	No (0)	No (0)
Q3: Comprehensive search performed?	Yes[a] (1)	Yes (1)
Q4: Publication type accounted for?	Yes[a] (1)	No (0)
Q5: Included/excluded studies mentioned?	Yes (1)	No (0)
Q6: Details of included studies reported?	Yes (1)	Yes (1)
Q7: Critical appraisal of scientific quality performed and described?	Yes[b] (1)	Yes (1)
Q8: Were the results of Q7 used in formulating the recommendations?	Yes (1)	Yes (1)
Q9: Appropriate analysis?	Yes[c] (1)	Yes (1)
Q10: Publication bias assessed?	No (0)	Yes (1)
Q11: Conflict of interest, overall and for included studies?	No (0)	No (0)
TOTAL[d]	8/11	7/11

Abbreviation: AMSTAR, Assessment of Multiple Systematic Reviews; NICE, National Institute of Clinical and Health Excellence; QUADAS, Quality Assessment of Diagnostic Accuracy Studies.
[a] Multiple databases searched. Explicit statements of searching grey/unpublished literature not made clear in relevant section but mentioned in overall guidance.
[b] QUADAS tool.
[c] Subgroup analysis by summary.
[d] Yes = 1 point; No = zero points.

Appendix 5. Search Strategy

Main Literature Search of Systematic Reviews and Primary Literature (2013-2015)

EBM Reviews – Cochrane Central Register of Controlled Trials February 2015, EBM Reviews – Cochrane Database of Systematic Reviews 2005 to February 2015, EBM Reviews – Database of Abstracts of Reviews of Effects 1st Quarter 2015, EBM Reviews – Health Technology Assessment 1st Quarter 2015, Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE® 1946 to Present, EMBASE 1974 to 2015 March 18 (Literature search update performed on February 17, 2016)

1 exp Prostatic Neoplasms/

2 (prostat$ adj3 (cancer$ or carcinoma$ or adeno$ or malignan$ or tumo?r$ or neoplas$)).mp.

3 1 or 2

4 exp Neoplasm Staging/

5 (staging or stage$1 or classif$ or evaluat$ or tnm).mp.

6 4 or 5

7 exp Magnetic Resonance Imaging/

8 Magnet$ resonance.mp.

9 (MRI or MR$2 or NMR$1).mp.

10 (MR adj (imag$ or scan$)).mp.

11 (magnet$ adj (imag$ or scan$)).mp.

12 ((magnet$ or MR) adj spectroscop$).mp.

13 or/7-12

14 (magnetic resonance imag$ or magnetic resonance spectroscop$).mp.

15 (dynamic adj4 (MRI or magnet$)).mp.

16 (diffusion weight$ adj3 (MRI or magnet$)).mp.

17 ((T1-weighted or T2-weighted or T3-weighted) adj3 imag$).mp.

18 (magnet$ adj (imag$ or spectrosop$ or scan$ or resonance)).mp.

19 (MPMRI or MP-MRI or MR$2 or DWI$ or DW-MRI or DCE$ or fmri).mp.

20 or/14-19

21 13 or 20

22 (comment or letter or editorial or note or erratum or short survey or news or newspaper article or patient education handout or case report or historical article).pt.

23 exp animal/ not humans/

24 22 or 23

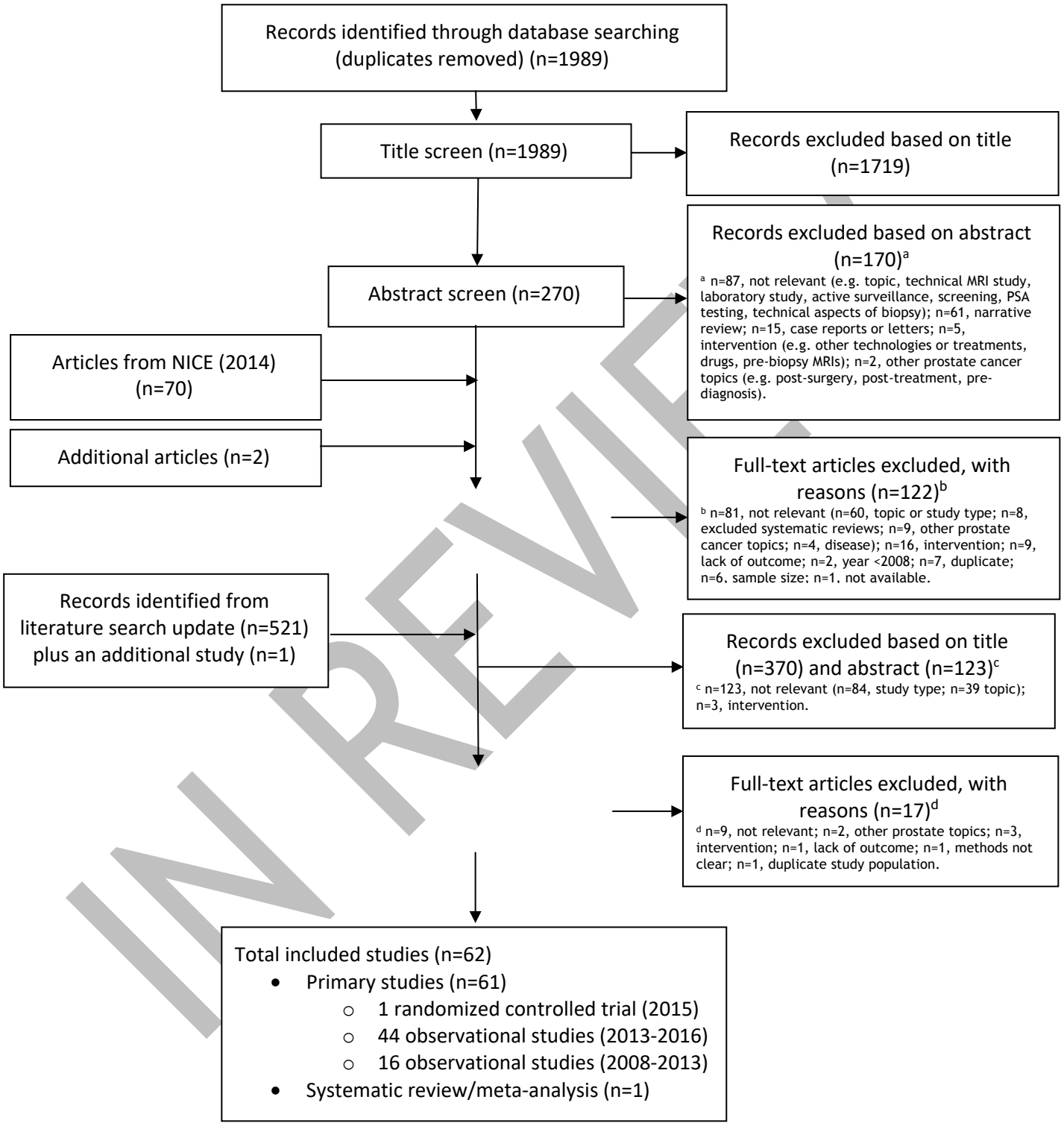
25 (3 and 6 and 21) not 24

26 limit 25 to (English language and yr="2013-Current") [Limit not valid in CDSR,DARE; records were retained]

27 remove duplicates from 26

Database of Abstracts of Reviews of Effects (DARE): a full-text database containing critical assessments of systematic reviews from a variety of medical journals produced by the National Health Services' Centre for Reviews and Dissemination (NHS CRD) at the University of York, England.
Health Technology Assessment (HTA): a database of abstracts that are descriptive in nature and are based on many different types of research including systematic reviews and ongoing and completed research based on trials, questionnaires and economic evaluations. Produced by the Centre for Reviews and Dissemination (CRD), and in collaboration with the INAHTA Secretariat (Sweden), the database comprises details on ongoing health technology assessments (studies of the medical, social, ethical, and economic implications of healthcare interventions) as well as publications reporting completed technology assessments carried out by organizations.

Appendix 6. Citation Flow Chart



Appendix 7. Summary Data and Tables

Table 1. Summary of Included Primary Studies (N=61)

Author (Year)	Study Location	PSA Risk Level[a]	Study Design	SS[b]
Orczyk et al (2016) [42]	United States, single-centre	Low	RCS/PCS	31
Bittencourt et al (2015) [43]	Netherlands, single-centre	Low	RCS	133
Boesen et al (2015) [44]	Denmark, single-centre	Intermediate	PCS	87
Counago et al (2015) [45]	Spain, single-centre	Low	RCS	274
de Cobelli et al (2015) [46]	Italy, single-centre	Low	RCS	223
Feng et al (2015) [47]	United States, single-centre	Low	RCS	112
Ghafoori et al (2015) [4]	Iran, single-centre	Intermediate	PCS	238
Junker et al (2015) [81]	Austria, single-centre	Low	PCS	50
Razi et al (2015) [48]	Iran, single-centre	Intermediate	PCS	80
Reisaeter et al (2015) [49]	Norway, single-centre	Low	RCS	63
Rud et al (2015) [41]	Norway, single-centre	Low	RCT	222/216
Schieda et al (2015) [3]	Canada, single-centre	Low	RCS	145
Woo et al (2015) [50]	Korea, single-centre	Intermediate	RCS	117
Counago et al (2014) [5]	Spain, single-centre	Low	RCS	150[c]
Gupta et al (2014) [6]	United States, single-centre	Low	RCS	60
Kan et al (2014) [7]	Hong Kong, single-centre	Low	RCS	56[d]
Kim et al (2014)(a) [8]	Korea, single-centre	Low	RCS	167
Kim et al (2014)(b) [9]	Korea, single-centre	Low	RCS	100
Kitamura et al (2014) [10]	Japan, single-centre	Low	PCS	54
Lawrence et al (2014) [11]	United Kingdom, single-centre	Low	PCS	40
Lista et al (2014) [12]	Spain, single-centre	Intermediate	PCS	85
Otto et al (2014) [13]	Germany, single-centre	Intermediate	RCS	37
Park et al (2014)(a) [14]	Korea, single-centre	Low	RCS	353
Park et al (2014)(b) [55]	Korea, single-centre	Low	RCS/PCS	282
Roethke et al (2014) [15]	Germany, single-centre	Low	RCS	376
Song et al (2014) [16]	Korea, single-centre	Low	RCS	382
Styles et al (2014) [17]	Australia, single-centre	-	PCS	38
Yao et al (2014) [75]	Japan, single-centre	Low	RCS	84
Armitage et al (2013) [18]	United Kingdom, single-centre	Intermediate	RCS/PCS	69[c]
Cerantola et al (2013) [19]	Switzerland, single-centre	Intermediate	RCS	60
Hegde et al (2013) [20]	United States, single-centre	Low	RCS	118
Hole et al (2013) [21]	Norway, single-centre	Intermediate	PCS	209
Isebaert et al (2013) [22]	Belgium, single-centre	Intermediate	PCS	75
Jeong et al (2013) [23]	Korea, single-centre	Intermediate	RCS/PCS	922
Johnston et al (2013) [24]	United Kingdom, single-centre	Low	RCS	568
Nepple et al (2013) [25]	United States, single-centre	Low	RCS	94
Pak et al (2013) [26]	Korea, single-centre	Intermediate	RCS	472[e]

Porcaro et al (2013) [27]	Italy, single-centre	Intermediate	RCS	154
Pugh et al (2013) [28]	United States, single-centre	Low	PCS	171
Renard-Penna et al (2013) [30]	France, single-centre	Low	RCS	101
Roethke et al (2013) [29]	Germany, single-centre	Low	RCS	385
Somford et al (2013) [31]	Netherlands, multi-centre	Intermediate	PCS	183
Soylu et al (2013) [32]	United States, single-centre	Intermediate	RCS	131
Tanaka et al (2013) [33]	Japan, single-centre	Low	PCS	67
Turkbey et al (2013) [34]	United States, single-centre	Low	RCS	133
Tsao et al (2013) [35]	Taiwan, single-centre	Intermediate	PCS	94
Guzzo et al (2012) [76]	United States, single-centre	Low	RCS	172
Jung et al (2012) [77]	United States, single-centre	Low	RCS	101
Kim et al (2012) [36]	Korea, single-centre	Intermediate	RCS	151
McClure et al (2012) [51]	United States, single-centre	Low	PCS	104
Panebianco et al (2012) [53]	Italy, single-centre	Low	PCS	105
Hwii Ko et al (2011) [78]	Korea, single-centre	Low	RCS	121
Novis et al (2011) [37]	Brazil, single-centre	Low	PCS	35
Ploussard et al (2011) [54]	France, multi-centre	Low	RCS/PCS	96
Turkbey et al (2011) [38]	United States, single-centre	Low	PCS	45
Lee et al (2010)(a) [79]	Korea, single-centre	Intermediate	RCS/PCS	67
Lee et al (2010)(b) [39]	Korea, single-centre	Intermediate	RCS	91
Turkbey et al (2010) [40]	United States, single-centre	Low	PCS	70
Brown et al (2009) [56]	United States, single-centre	Low	RCS	62
Zhang et al (2009) [80]	United States, single-centre	Low	RCS	158
Cirillo et al (2008) [57]	Italy, single-centre	Low	RCS	143

Abbreviations: PCS, prospective case series; PSA, prostate-specific antigen; RCS, retrospective case series; RCT, randomized controlled trial; SS, sample size.
[a] Risk of spread according to mean or median PSA reported in study population: low (<10 ng/mL), intermediate (10-20 ng/mL), high (>20 ng/mL), or PSA category with the highest proportion of patients.
[b] After exclusions; intervention/control.
[c] Only a subgroup of 47 patients in Counago et al, 2014 and only a subgroup of 35 patients in Armitage et al, 2013 had magnetic resonance imaging and radical prostatectomy.
[d] Eligible scans.
[e] Analysis performed by lobes.

Table 2. Detailed Summary of Included Primary Studies (N=61)

Study (Year)	Patient Enrollment	Study Population	CS	Bp	Core No.	Bp to MRI	Age and PSA Results of Study Population	Pathological Staging Results [no. (%)]
Orczyk et al (2016) [42]	Pts scheduled for RP+pre-op MRI, 2011-2012	Inclusion: pre-op MRI	Bp	Y	NR	NR	Mean age: 61 yrs (SD: 9.0); mean PSA: 5.8 ng/mL (SD: 2.7)	pT0: 1 (3.2), pT2a: 1 (3.2), pT2b: 0 (0), pT2c: 18 (58.1), pT3a: 11 (35.5)
Bittencourt et al (2015) [43]	Consecutive pts to referral centre, 2010-2013	Inclusion: RP, pre-op MRI, prospective PI-RADS classification Exclusion: prior RT or HT, >6 mo. to RP after MRI	Bp	Y	NR	NR	Mean age: 61.0 (R: 47.0-72.0); median PSA: 8.6 ng/mL (R: 2.5-76.0)	pT2a: 11 (8.3), pT2b: 1 (0.8), pT2c: 61 (45.9), pT3a: 44 (33.1), pT3b: 11 (8.3), pT4: 5 (3.8)
Boesen et al (2015) [44]	Pts, 2011-2013	Inclusion: dx PrCa (localized) by DRE+TRUS, RP scheduled, no prior trt Exclusion: contraindications to MRI	DRE, TRUS	Y	NR	NR	Median age: 65.0 (R: 47.0-74.0); median PSA: 11.0 ng/mL (4.6-45.0)	≤pT2: 55 (63.2); pT3a: 32 (36.8) [31 pts for EPE, 5 pts for EPE+SVI, 1 pt for SVI only]
Counago et al (2015) [45]	Pts, 2009-2015	Inclusion: histological dx PrCa, MRI, trt by RT±HT or RP Exclusion: trt prior to MRI, no MRI or CS	DRE, TRUS	Y	NR	NR	Mean age: 66.8 (SD: 8.1); PSA <10 ng/mL: 179 (65.3), PSA 10-20 ng/mL: 66 (24.1), PSA >20 ng/mL: 29 (10.6)	pT1-T2a: 23 (25.6), pT2b-T2c: 57 (63.3), pT3-T4: 10 (11.1)
de Cobelli et al (2015) [46]	Pts who had RP, 2009-2014	Inclusion: pts fulfilled criteria for AS, no HT, had RP	DRE, PSA, TRUS	Y	12	≥6-8 w	Mean age: 62.3 (SD: 8.3); mean PSA: 6.0 ng/mL (SD: 1.9)	pT2a: 23 (10.3), pT2b: 3 (1.4), pT2c: 145 (65.0), pT3a: 45 (20.2), pT3b: 7 (3.1)
Feng et al (2015)[a] [47]	Consecutive pts, 2010-2013	Inclusion: had MRI and RP Exclusion: prior HT or RT, MRI prior to Bp	DRE	Y	NR	4 w	Mean age: 62.8 (SD: 7.5); mean PSA: 8.2 ng/mL (SD: 7.2)	pT2a/b: 8 (7.1), pT2c: 70 (62.5), pT3a: 15 (13.4), pT3b: 19 (17.0)
Ghafoori et al (2015) [4]	Pts with proven PrCa, 2011-2013	Inclusion: pts proven PrCa by TRUS	TRUS	Y	NR	≥4 w	Mean age: 67.4 yrs (SD: 9.0); mean PSA: 16.7 ng/mL (SD: 18.0)	pT3b: 63 (23.5), otherwise NR
Junker et al (2015) [81]	Consecutive pts, 2012-2013	Inclusion: Bp PrCa, RP, MRI prior to surgery Exclusion: contraindications, poor pathology results	Bp	Y	NR	≥3 m	Mean age: 63.0 yrs (SD: 8.0); mean PSA: 7.3 ng/mL (SD: 4.3)	NR
Razi et al (2015) [48]	Consecutive pts, 2009-2012	Inclusion: clinically localized PrCa, had ER+MRS prior to RP	DRE, PSA, TRUS	Y	NR	2 w	Mean age: 63.5 yrs (SD: 7.7); mean PSA: 16.3 ng/mL (SD: 19.0)	pT2: 45 (56.3) pT3: 25 (31.3), pT4: 8 (10.0); benign tumour is 2 (2.5)
Reisaeter et al (2015) [49]	Consecutive pts, 2010	Inclusion: Bp-proven PrCa, MRI prior to RP Exclusion: insignificant PrCa, no data	Bp	Y	NR	14 w	Median age: 61.6 yrs (R: 42.9-70.3); median PSA: 8.0 ng/mL (R: 3.0-81.4)	NR
Rud et al (2015) [41]	Pts scheduled for RB, 2009-2012	Exclusion: prior prostate MRI, MRI contraindications, hip prosthesis	DRE, TRUS	Y	NR	11 w	MRI+DWI, mean age: 62.0 yrs (SD: 6.0); median PSA: 7.8 ng/mL (IQR: 4.9-11.3); non-MRI: mean age: 63.0 yrs (SD: 6.0); median PSA: 8.2 ng/mL (IQR: 6.2-18.3)	MRI+DWI, pT2: 103 (46.4), pT3 119 (53.6) Non-intervention: pT2: 111 (51.4), pT3 105 (48.6)
Schieda et al (2015) [3]	Consecutive pts, 2012-2014	Inclusion: RP, pre-op MRI with 3 T, no prior trt	PSA, Bp	Y	NR	NR	Mean age: 62.8 yrs (SD: 6.0); mean PSA: 9.0 ng/mL (SD: 8.6)	pT3a: 95 (65.5), otherwise NR

Study (Year)	Patient Enrollment	Study Population	CS	Bp	Core No.	Bp to MRI	Age and PSA Results of Study Population	Pathological Staging Results [no. (%)]
		Exclusion: lack of RP data, pre-trt						
Woo et al (2015) [50]	Database pts, 2013	Inclusion: had MRI and then RP Exclusion: prior trt, Bp or MRI	PSA, Bp	Y	NR	NR	Mean age: 68.0 yrs (SD: 6.8); mean PSA: 12.4 ng/mL (SD: 13.1)	pT3a: 50 (42.7), otherwise NR
Counago et al (2014) [3]	Chart review, 2009-2013	Inclusion: histological dx of PrCa, MRI+body coil before trt (RP or RT)	DRE, TRUS	Y	NR	>3 w	Mean age: 61.0 yrs (SD: 6.8); PSA <10 ng/mL: 34 (72.3), PSA 10-20 ng/mL: 10 (21.1), PSA >20 ng/mL: 1 (2.1), unkwn: 2 (4.2)	pT2a: 5 (10.6), pT2b: 5 (10.6), pT2c: 30 (63.8), pT3a: 4 (8.5), pT3b: 3 (6.4)
Gupta et al (2014)[6]	Pts undergoing MRI for PrCa, 2011	Inclusion: complete MRI then RP Exclusion: incomplete MRI, pre-trt, ER unable.	DRE	Y	NR	13 w	Mean age: 60.1 yrs (SD: 7.3); mean PSA: 6.9 ng/mL (R: 1.2-46.3)	≤pT2: 38 (63.3); pT3a: 18 (30.0); pT3b: 4 (6.7)
Kan et al (2014) [7]	Consecutive pts of RP, 2010-2012	Inclusion: 12-core Bp+pathology reports.	TRUS	Y	≥ 12	NR	Age: NR; PSA <10 ng/mL: 64 (64.0), PSA 10-20: 29 (29.0), PSA >20: 7 (7.0)	NR
Kim et al (2014)(a) [8]	Referrals, 2006-2008	Inclusion: Bp-proven PrCa, no hx trt, MRI prior to surgery.	Bp	Y	NR	3 w	Mean age: 66.5 yrs (R: 52.0-78.0); mean PSA: 8.5 ng/mL (R: 1.1-37.3)	pT3a: 23 (13.8); pT3b: 7 (4.2)
Kim et al (2014)(b) [9]	Hospital medical data, 2008-2012	Inclusion: dx by 12-core TRUS Bp, localized, core length ≤7 mm, total cancer length ≤10 mm, low D'Amico risk, pre-trt MRI, negative bone scans.	TRUS	Y	12	≥6 w	Median age: 63.3 yrs (R: 51.0-76.0); median PSA: 6.5 ng/mL (R: 2.2-9.5)	pT2a: 28 (28.0), pT2b: 23 (23.0), pT2c: 36 (36.0), pT3a: 10 (10.0), pT3b: 3 (3.0)
Kitamura et al (2014) [10]	Consecutive pts, 2009-2013	Inclusion: cT1c-T2N0M0 Bp-proven localized PrCa, MRI plus TRUS, followed by RP, no hx trt	Bp	Y	NR	NR	Mean age: 62.7 yrs (SD: 6.4); median PSA: 5.7 ng/mL (4.4-7.6)	pT2a: 17 (31.5), pT2b: 2 (3.7), pT2c: 24 (44.4), pT3a: 7 (13.0), pT3b: 4 (7.4)
Lawrence et al (2014) [11]	PrCa pts, 2010-2014	Inclusion: Bp-proven PrCa of intermediate to high-risk as defined, plan for RP, Bp ≥6 wks prior to MRI	DRE	Y	NR	≥6 w	Median age: 62.5 yrs (R: 42.0-73.0); median PSA: 7.3 ng/mL (R:0.6-14.6)	≤pT2: 17 (42.5); pT3a: 23 (57.5)
Lista et al (2014) [12]	Prca pts	Inclusion: proximity of PrCa to NVB in pts for NS RP, need for extended pre-op planning in high-risk pts	DRE, TRE	Y	NR	4 w	Mean age: 63.7 yrs (SD: 6.9); mean PSA: 12.6 ng/mL (SD: 13.8)	pT2a: 12 (14.1), pT2b: 3 (3.5), pT2c: 37 (43.5), pT3a: 19 (22.3), pT3b: 14 (16.5)
Otto et al (2014) [13]	Consecutive pts, 2010	Inclusion: histologically dx PrCa Exclusion: contraindications to MRI	PSA, TRUS	Y	μ: 11.6	8 w	Mean age: 65.0 yrs (R: 53.0-75.0); median PSA: 13.5 ng/mL (R: 3.7-56.0)	pT3a: 10 (27.0); pT3b: 5 (13.5)
Park et al (2014)(a) [14]	Pts who had RB, 2008-2011	Inclusion: Bp-proven PrCa Exclusion: hx trt (HT or RT), MRI performed at other institutions	Bp, DRE, PSA	Y	NR	≥3 w	Median age: 64.0 yrs (R: 43.0-73.0); median PSA: 5.3 ng/mL (R: 1.7-58.5)	NR
Park et al (2014)(b) [55]	Consecutive pts with PrCa+RP, 2007-2009	Exclusion: hx trt (HT or RT), poor imaging quality, missing clinical data, lack of FU data	DRE	Y	NR	4 w	Median age: 64.0 yrs (R: 38.0-88.0); median PSA: 6.6 ng/mL (R: 0.3-57.0)	pT3a: 68 (24.0); pT3b: 22 (8.0); pLNM: 2 (0.7)
Roethke et al (2014) [15]	Referrals, 2004-2008	Inclusion: Bp-proven PrCa scheduled for RP Exclusion: hx trt (HT or RT)	Bp	Y	NR	≥6 w	Median age: 63.0 yrs (R: 43.0-6.0); median PSA: 8.8 ng/mL (R: 1.1-2.5)	pT3b: 35 (9.3)

Study (Year)	Patient Enrollment	Study Population	CS	Bp	Core No.	Bp to MRI	Age and PSA Results of Study Population	Pathological Staging Results [no. (%)]
Song et al (2014)[b] [16]	MRR, 2007-2012	Inclusion: 12-core Bp, pre-op MRI, RP, low risk PrCa by D'Amico	Bp	Y	12	NR	Mean age: 64.1yrs (SD: 7.6); mean PSA: 5.1 ng/mL (SD: 1.8)	pT2: 330 (86.4), pT3a: 49 (12.8), pT3b: 2 (0.5)
Styles et al (2014) [17]	Pts, 2010-2012	Inclusion: Bp-proven PrCa, localized based on clinical staging, undergo RP post-MRI	DRE	Y	NR	3-8 w	Mean age: 60.0 yrs (R: 42-73); PSA: NR	NR
Yao et al (2014) [75]	Consecutive pts, 2010-2012	Inclusion: dx PrCa by TRUS, RP Exclusion: hx HT	PSA, TRUS	Y	8-14	>4 w	Mean age: 64.0 yrs (SD: 7.0); mean PSA: 9.3 ng/mL (R: 2.7-34.6)	pT2: 59 (70.2); pT3a: 21 (25.0); pT3b: 4 (4.8)
Armitage et al (2013) [18]	Database review, all RP pts, 2008	Inclusion: all pts who had a RP	CS	Y	NR	6 w	Mean age: 64.1 yrs (R: 46-74); PSA ≤10 ng/mL: 4 (5.8), PSA 10-20 ng/mL: 58 (84.1), PSA >20 ng/mL: 7 (10.1)	≥pT3: 22/35 (62.9)
Cerantola et al (2013) [19]	All RP pts, 2008-2012	Inclusion: MRI performed in single-centre	DRE, PSA, TRUS	Y	12	≥3 w	Mean age: 67.0 yrs (SD: 7.0); mean PSA: 12.7 ng/mL (SD: 12.7)	pT2a: 3 (5), pT2b: 1 (1.7), pT2c: 25 (41.7), pT3a: 22 (36.7), pT3b: 7 (11.7), pT4: 2 (3.3)
Hegde et al (2013) [20]	RP pts+pre-op 3T MRI, 2008-2011	Inclusion: no hx trt, dx cT1-cT2 PrCa and suspicion of T3 disease	PSA, TRUS	Y	≥12 (86% of men)	6 w	~Age: 58.7 yrs (R: 52.4-65.2); ~PSA: 6.5 ng/mL (R: 3.1-10.3)	pT2a: 10 (8.5), pT2b: 3 (2.5), pT2c: 76 (64.4), pT3a: 19 (16.1), pT3b: 10 (8.5)
Hole et al (2013) [21]	Consecutive pts, pre-op MRI+RP, 2007-2010	Inclusion: PrCa pts, pre-op MRI, RP Exclusion: no hx trt, no skeletal metastases	DRE, TRUS	Y	NR	13 w	Mean age: 62.5 yrs (SD: 5.9); mean PSA: 19.8 ng/mL (SD: 29.4)	pT2a: 11 (5.3), pT2b: 0 (0), pT2c: 62 (29.7), pT3a: 96 (45.9), pT3b: 34 (16.3); pT4: 5 (2.4); Tx: 1 (0.5)
Isebaert et al (2013) [22]	Consecutive pts, 2008-2011	Inclusion: newly dx Bp-proven PrCa, intermediate to high-risk PrCa by D'Amico, no PLN by CE-CT and LNM risk ≥10% and ≤35% by Partin tables, negative bone scan	PSA, TRUS	Y	NR	≥4 w	Median age: 66.0 yrs (R: 49.0-74.0); median PSA: 10.4 ng/mL (R: 1.5-70.9)	pT2b: 2 (2.7), pT2c: 31 (41.3), pT3a: 20 (26.7), pT3b: 19 (25.3); pT4: 3 (4.0)
Jeong et al (2013) [23]	Consecutive pts, PrCa+RP+PL, 2000-2012	Inclusion: MRI, high-risk, treated RP Exclusion: hx trt	Bp	Y	NR	4 w	Mean age: 66.1 yrs (SD: 6.7); medan PSA: 10.7 ng/mL (R: 0.3-737.0)	≤pT2: 387 (41.9); pT3a: 530 (57.5); pT3b: 117 (12.7); pLN: 58 (6.3)
Johnston et al (2013) [24]	Database review, all pts, RB-RP+pre-op MRI, 2005-2011	Inclusion: pts undergoing RB-RP and had a pre-op MRI	CS	Y	NR	4 w	Median age: 62.0 yrs (R: 35.0-74.0); median PSA: 8.7 ng/mL (R:0.5-63.0)	pT3a: 280 (49.3), pT3b: 34 (6.0)
Nepple et al (2013)[c] [25]	Consecutive pts, RP+pre-op MRI, 2003-2008	Inclusion: PrCa with risk factors for local extension (gleason grade ≥4 +3, PSA ≥10 ng/mL, abnormal DRE, or extensive Bp involvement)	Bp, DRE, PSA	Y	NR	6 w	Median age: 61.0 yrs (48.0-72.0); median PSA: 7.0 ng/mL (R: 1.3-35.0)	pT3a: 22 (24.2); pT3b: 8 (8.8)
Pak et al (2013) [26]	Referral pts, RP, 2007-2012	Inclusion: undergoing RP Exclusion: hx trt, pts without detailed information	DRE, PSA, TRUS	Y	12	NR	~Age: 65.3 yrs; ~PSA: 11.7 ng/mL	pT3a: 194 (20.6)

Study (Year)	Patient Enrollment	Study Population	CS	Bp	Core No.	Bp to MRI	Age and PSA Results of Study Population	Pathological Staging Results [no. (%)]
Porcaro (2013) [27]	Pts pre-op MRI, 2003-2006	Exclusion: pts using 5-alpha reductase inhibitors, LH-releasing hormone analog, anti-androgens or testosterone replacement treatment	Bp, DRE PSA	Y	NR	NR	Mean age: 66.0 yrs (SD: 6.0); mean PSA: 11.0 ng/mL (SD: 11.4)	pT2: 97 (63.0), pT3a: 41 (26.6), pT3b: 16 (10.4)
Pugh et al (2013) [28]	Consecutive pts, 2006-2010, RB RP	Inclusion: T1c-T2c, pre-trt PSA <10 ng/mL, GS of 7, and pre-surgery MRI	PSA, TRUS	Y	NR	NR	Median age: 60.0 yrs (R: 42.0-76.0); median PSA: 4.9 ng/mL (R: 0.4-9.9)	pT2a: 114 (66.7), pT3a: 38 (22.2), pT3b: 19 (11.1)
Renard-Penna et al (2013) [30]	Consecutive pts, pre-op MRI + RP, 2009-2010	Inclusion: Bp-proven PrCa, RP planned within 1 mo. after MRI, ≥8 w between biopsy and MRI, able pts for MRI, no hx HT	DRE	Y	NR	≥8 w	Median age: 60.0 yrs (R: 39.0-71.0); mean PSA: 8.0 ng/mL (SD: 4.4)	pT2a: 8 (7.9), pT2b: 10 (9.9), pT2c: 67 (66.3), pT3a: 10 (9.9), pT3b: 6 (5.9)
Roethke et al (2013) [29]	Pts referrals, 2003-2008	Inclusion: Bp-proven PrCa, pre-op MRI, RP Exclusion: hx HT or RT	Bp, PSA	Y	NR	≥6 w	Mean age: 62.7 yrs (R: 42.0-77.0); mean PSA: 8.9 ng/mL (R: 0.4-52.5)	pT2: 268 (69.6), ≥ pT3: 117 (30.4)
Somford et al (2013) [31]	Consecutive pts, RP, 2007-2010	Exclusion: pts with extensive T3 PrCa not considered for RP, nodal metastasis on MRI	DRE	Y	8-12 cores	NR	Mean age: 62.4 yrs (SD: 4.9); mean PSA: 10 ng/mL (SD: 8.4)	pT2: 92 (50.3), pT3: 91 (49.7) [pT3b: 21 (11.5%)]
Soylu et al (2013) [32]	Consecutive pts, pre-op MRI+RP, 2007-2010	Exclusion: hx trt (HT, RT, CT), incomplete MRI examination	Bp	Y	NR	NR	Median age: 68.0 yrs (R: 43.0-75.0); median PSA: 12.1 ng/mL (R: 1.5-65.0)	pT3b: 23 (17.6)
Tanaka et al (2013) [33]	Pts, T2 or T3 on MRI, 2010-2012	Exclusion: no hx HT	Bp	Y	12	≥8 w	Median age: 67.0 yrs (R: 51.0-74.0); median PSA: 7.0 ng/mL (R: 2.87-27.6)	pT2a: 11 (16.4), pT2b: 6 (9.0), pT2c: 33 (49.3), pT3a: 15 (22.4), pT3b: 2 (3.0)
Turkbey et al (2013)[d] [34]	Pts, 2007-2010	Inclusion: clinical-pathologic parameters available to calculate AS eligibility by established risk criteria, pre-op mpMRI at 3.0 T followed by RB	DRE PSA, Bp	Y	NR	NR	Median age: 59.0 yrs (R: 39.0-74.0); median PSA: 4.4 ng/mL (R: 0.9-48.9)	pT3a: 46 (34.6); pT3b: 6 (4.5)
Tsao et al (2013) [35]	Dx PrCa pts, 2001-2007	Inclusion: pre-op MRI, RP Exclusion: hx HT or RT	DRE or PSA	Y	10-12	>6 w	Mean age: 68.9 yrs (R: 50.0-85.0); mean PSA: 16.9 ng/mL (R: 0.1-107.0)	pT2a: 65 (69.1), pT3a: 12 (12.8), pT3b: 17 (18.1)
Guzzo et al (2012)[e] [76]	Database review, consecutive RP pts, 1991-2007	Inclusion: pre-op MRI+clinical criteria for AS (Johns Hopkins expectant management program)[e]	Bp, PSA	Y	NR	NR	Mean age: 59.8 yrs (SD: 6.2); mean PSA: 5.2 ng/mL (SD: 2.2)	pT2a: 105 (76.1), pT2b: 3 (2.2), pT2c: 17 (12.3), pT3a: 11 (8.0), pT3b: 2 (1.4)
Jung et al (2012) [77]	Database review, consecutive pts, 1997-2009	Inclusion: Bp-proven PrCa, pre-op MRI, RP Exclusion: hx neoadjuvant therapy	Bp, PSA	Y	NR	≥6 w	Mean age: 59.0 yrs (R: 42.0-75.0); median PSA: 6.9 ng/mL (R: 1.1-38.0)	pT3: 23 (22.7); otherwise NR
Kim et al (2012) [36]	Consecutive pts, 2005-2010	Inclusion: Bp-proven PrCa, RP	Bp, PSA	Y	NR	4 w	~Age: 65.8 yrs (R: 47.0-76.0); ~PSA: 12.0 ng/mL (R: 22.0-45.0)	pT3a: 81 (53.6); pT3b: 34 (22.5); otherwise NR

Study (Year)	Patient Enrollment	Study Population	CS	Bp	Core No.	Bp to MRI	Age and PSA Results of Study Population	Pathological Staging Results [no. (%)]
		Exclusion: contraindications for MRI, hx neoadjuvant HT or RT, <3 w between Bp and MRI						
McClure et al (2012) [51]	Consecutive pts, 2004-2008	Inclusion: had MRI + ER prior to RB Exclusion: open RP	Bp, PSA	Y	NR	13 w	Mean age: 60.1 yrs (SD: 6.8); mean PSA: 6.5 ng/mL (SD: 4.7)	pT2: 198 (95.2); pT3: 10 (4.8)
Panebianco et al (2012) [53]	Consecutive pts, 2006-2010	Inclusion: Bp-proven PrCa, BL/NS/RP based on clinical assessment [T1c or T2a, PSA <10 ng/mL, GS <8, only 1 Bp GS >6 at ipsilateral side Exclusion: hx HT/RT/CT, prior prostate surgery, contraindications for MRI	DRE, PSA, TRUS	Y	12	≥6 w	Median age: 57.0 yrs (R: 48.0-66.0); ~median PSA: 5.5 ng/mL: (R: 2.5-9.8)	pT2a: 36 (34.3), pT2b: 58 (55.2), pT3a: 11 (10.5), pT3b: 0 (0)
Hwii Ko et al (2011) [78]	Pts, 2007-2009	Inclusion: clinical stage T1c to T3c, dx PrCa by TRUS, RP, pre-op MRI Exclusion: hx HT, RT or aby ablative technique	Bp, DRE	Y	NR	4 w	Mean age: 62.8 yrs (R: 46.0-74.0); mean PSA: 9.6 ng/mL (R: 0.4-24.4)	pT2a: 17 (14.0), pT2b: 11 (9.1), pT2c: 50 (41.3), pT3a: 21 (17.4), pT3b: 7 (5.8); pT3c: 10 (8.3); pT4: 5 (4.1)
Novis et al (2011) [37]	Referral pts, 2005-2006	Inclusion: US-guided Bp-proven PrCa, no hx HT, RP candidate, T1c-T2a, GS ≤6, PSA ≤10 ng/mL	Bp, DRE, PSA	Y	NR	≥3 w	Mean age: 64.9 yrs (R: 50.0-77.0); mean PSA: 6.1 ng/mL (R: 2.6-10.0)	pT2a: 2 (5.7), pT2b: 3 (8.6), pT2c: 11 (31.4), pT3a: 8 (22.9), pT3b: 3 (8.6)
Ploussard et al (2011)[f] [54]	Pts, 21-core Bp for PSA >0.7 ng/mL/yr, ab DRE, PSA >4 ng/mL and/or PSA-R <10%, 2001-2008	Inclusion: PSA ≤10 ng/mL, T1-T2a, GS ≤6, >10 yrs life expectancy (<65 yrs w/o comorbidities, >65 yrs w/o oncogeriatrics medical visit) Exclusion: tumour ≥3 cores, length/core ≥3 mm, no RP, no pre-op MRI	Bp, DRE, PSA	Y	21	≥6 w	Mean age: 62.4 yrs (R: 51.0-73.2); mean PSA: 6.1 ng/mL (R: 1.8-10.0)	pT2a: 26 (27.1), pT2b: 2 (2.1), pT2c: 50 (52.1), pT3: 17 (17.7) [EPE: 16 (16.7), SVI: 1 (1.0)]
Turkbey et al (2011) [38]	Consecutive pts, 2008-2009	Inclusion: Bp-proven PrCa, RB within 180 days of pre-op MRI w/o any treatment Exclusion: contraindications to MRI or ER	Bp, PSA	Y	NR	≥10 w	Median age: 60.0 yrs (R: 49.0-75.0); median PSA: 5.8 ng/mL (R: 2.3-23.7)	pT3a: 12 (26.7); pT3b: 2 (4.4)
Lee et al (2010)(a) [79]	Pts, 1998-2006	Inclusion: cT3 PrCa, RPP, pre-op MRI Exclusion: hx HT or RT	DRE, PSA, TRUS	Y	6-10	>3-4 w	~Age: 64.5 yrs; ~PSA: 13.7 ng/mL (2.7-207.0)	≤pT2c: 53 (79.1), pT3a: 3 (4.5), pT3b: 11 (16.4)
Lee et al (2010)(b) [39]	Record review, consecutive pts, 2007-2009	Inclusion: pre-op 1.5 T MRI with ER or pelvic coil, RP Exclusion: positive nodes, bone metastasis, technical difficulties with imaging or pathology	Bp, PSA	Y	~11.8	3 w	Mean age: 63.0 yrs (SD: 1.5); mean PSA: 10.7 ng/mL (SD: 1.9)	pT2: 34 (37.4); pT3a: 49 (53.8); pT3b: 8 (8.8)
Turkbey et al (2010) [40]	Pts, 2004-2007	Inclusion: TRUS Bp-proven PrCa, RP performed within 180 days of imaging and no trt	Bp, PSA	Y	NR	12 w	Mean age: 60.4 yrs (R: 40.0-75.0); mean PSA: 5.5 ng/mL (1.0-19.9)	pT2: 47 (67.1); pT3: 23 (32.9)

Study (Year)	Patient Enrollment	Study Population	CS	Bp	Core No.	Bp to MRI	Age and PSA Results of Study Population	Pathological Staging Results [no. (%)]
		Exclusion: contraindication to MRI, unable for ER						
Brown et al (2009) [56]	Pts, RP+pre-op MRI, 2002-2005	Inclusion: pts able and willing to undergo MRI, risk of ECE based on pre-op assessment	Bp, PSA	Y	NR	≥8-10 w	Mean age: 58.0 yrs (R: 42.0-73.0); median PSA: 6.1 ng/mL (R: 0.5-30.5)	pT2: 41 (66.1), pT3: 21 (33.9)
Zhang et al (2009) [80]	Consecutive pts, 2003-2004	Inclusion: 1.5 T MRI+post-RP in time period, T1c PrCa, pathology available Exclusion: hx HT/CT/RT	Bp, PSA	Y	NR	10 w	Median age: 58.0 yrs (R: 40.0-76.0); median PSA: 5.3 ng/mL (R:1.5-21.0)	pT2: 124 (78.5), pT3a: 29 (18.4), pT3b: 2 (1.3), pT4: 2 (1.3) [No tumour: 1 (0.6)]
Cirillo et al (2008) [57]	Pts, 2002-2005	Inclusion: Bp-proven PrCa+pre-op MRI	Bp, DRE, PSA	Y	NR	NR	Mean age: 70 yrs (R: 52.0-83.0); mean PSA: 9.1 ng/mL (R: 1.6-136.5)	NR

Abbreviations: ab, abnormal; Bp, biopsy; CE-CT, contrast-enhanced computed tomography; CS, clinical staging; CT, chemotherapy; d, days; DCE, dynamic contrast-enhanced imaging; DRE, digital rectal examination; DWI, diffusion-weighted imaging; dx, diagnosis; EPE, extraprostatic extension; ER, endorectal coil; grps, groups; HT, hormone therapy; hx, history; IQR, interquartile range; LN, lymph node; LNM, lymph node metastasis; M, median; mo, months; μ, mean; MRI; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRR, medical record review; MRS, magnetic resonance spectroscopic imaging; pts, patients; NR, not reported; OCD, organ-confined disease; p, pathology; PL, pelvic lymphadenopathy; PLN, pelvic lymph nodes; PrCa, prostate cancer, pre-op, preoperative; PSA, prostate-specific antigen; PSA-R, free-to-total PSA ratio; R, range; RB, robotic-assisted surgery; RP, radical prostatectomy; RT, radiation therapy; SD, standard deviation; SVI, seminal vesicle invasion; T, Tesla; trt, treatment; TRE, transrectal echography; TRUS, transrectal ultrasound; US, ultrasound; w, weeks; w/o, without.

[a] Additional information taken from prior Feng et al, 2015 study [88].

[b] Pathology, for n=381 patients.

[c] Of 91 patients with final pathology data.

[d] Remainder of pathology stage data not presented, therefore presented data does not add up to the total sample size of 133.

[e] Pathology, for n=138 patients.

[f] Clinical criteria includes PSA density ≤0.15 ng/mL/cm^3, no more than 2 biopsy cores positive for cancer, no more than 50% of 1 core involved with cancer, and no Gleason pattern 4 or 5 identified on biopsy.

[g] Percentages do not add up to 100% and sample size only to 95, based on original paper.

Note 1. The mean and corresponding SD or median and corresponding ranges were abstracted when available. Otherwise, the mean and range were abstracted where available, or mean alone. When the original paper provided age or PSA among subgroups, the average was calculated. This is indicated in tables as ~. Where ranges were quite extreme, the median was presented in lieu of the mean.
Note 2. OCD translated to ≤pT2, EPE translated to pT3a, and SVI translated to pT3b.

Table 3. Summary: Use of MRI ± DCE, DWI, MRS (N=61 Studies)

Author (Year)	Experienced Reader	Complete/ Partially Blinded	1.5 T MRI	3 T MRI	ER	BC/PC	+DCE	+DWI	+MRS
Randomized Controlled Trials (n=1)									
Rud et al (2015) [41]	∫		∫			∫		∫	
Observational Studies (n=66)									
Orczyk et al (2016) [42]	∫	∫		∫		∫	∫	∫	
Bittencourt et al (2015) [43]	∫			∫	∫[a]	∫[a]	∫	∫	
Boesen et al (2015) [44]	∫	∫		∫		∫	∫	∫	
Counago et al (2015) [45]	∫	∫		∫			∫	∫	
de Cobelli et al (2015) [46]			∫		∫		∫	∫	
Feng et al (2015) [47]	∫	∫		∫		∫	∫	∫	
Ghafoori et al (2015) [4]	∫		∫		∫	∫			
Junker et al (2015) [81]	∫	∫		∫		∫	∫	∫	
Razi et al (2015) [48]	∫	∫	∫		∫				∫
Reisaeter et al (2015) [49]	∫	∫	∫		∫	∫	∫	∫	
Schieda et al (2015) [3]	∫			∫		∫	∫	∫	
Woo et al (2015) [50]	∫	∫		∫				∫	
Counago et al (2014) [5]	∫			∫		∫	∫	∫	
Gupta et al (2014) [6]	∫	∫		∫	∫	∫	∫	∫	
Kan et al (2014) [7]	∫		∫			∫		∫	
Kim et al (2014)(a) [8]	∫	∫		∫		∫		∫	
Kim et al (2014)(b) [9]	∫	∫		∫		∫	∫	∫	
Kitamura et al (2014) [10]	∫	∫	∫		∫	∫		∫	∫
Lawrence et al (2014) [11]	∫	∫		∫		∫		∫	
Lista et al (2014) [12]			∫		∫	∫	∫	∫	
Otto et al (2014) [13]	∫	∫		∫	∫	∫	∫	∫	∫
Park et al (2014)(a) [14]	∫	∫		∫		∫	∫	∫	
Park et al (2014)(b) [55]	∫	∫		∫		∫	∫	∫	
Roethke et al (2014) [15]	∫		∫		∫	∫			
Song et al (2014) [16]	∫			∫		∫		∫	
Styles et al (2014) [17]		∫		∫	∫		∫	∫	∫
Yao et al (2014) [75]			∫			∫		∫	
Armitage et al (2013) [18]	∫		∫			∫			
Cerantola et al (2013) [19]	∫			∫	∫		∫	∫	
Hegde et al (2013) [20]	∫			∫	∫	∫	∫	∫	
Hole et al (2013) [21]	∫		∫			∫		∫	
Isebaert et al (2013) [22]	∫	∫	∫			∫	∫	∫	
Jeong et al (2013) [23]	∫		∫	∫	∫	∫		∫	
Johnston et al (2013) [24]	∫		∫						
Nepple et al (2013) [25]	∫		∫		∫				
Pak et al (2013) [26]	∫			∫				∫	

Porcaro et al (2013) [27]	✓	✓	✓		✓	✓					
Pugh et al (2013) [28]	✓		✓		✓	✓					
Renard-Penna et al (2013) [30]	✓	✓	✓			✓	✓				
Roethke et al (2013) [29]	✓		✓		✓	✓					
Somford et al (2013) [31]	✓	✓		✓	✓	✓	✓	✓			
Soylu et al (2013) [32]	✓	✓	✓		✓	✓	✓	✓			
Tanaka et al (2013) [33]				✓		✓	✓	✓			
Turkbey et al (2013) [34]	✓	✓		✓	✓	✓[b]	✓	✓	✓		
Tsao et al (2013) [35]	✓		✓				✓				
Guzzo et al (2012) [76]	✓	✓	✓		✓	✓[c]					
Jung et al (2012) [77]	✓	✓	✓		✓	✓					
Kim et al (2012) [36]	✓			✓	✓	✓					
McClure et al (2012) [51]	✓	✓	✓		✓		✓	✓	✓		
Panebianco et al (2012) [53]	✓	✓		✓	✓	✓	✓	✓	✓		
Hwii Ko et al (2011) [78]	✓	✓		✓		✓					
Novis et al (2011) [37]	✓	✓	✓		✓	✓	✓		✓		
Ploussard et al (2011) [54]	✓	✓	✓		✓						
Turkbey et al (2011) [38]	✓	✓		✓	✓	✓[c]	✓	✓	✓		
Lee et al (2010)(a) [79]	✓			✓		✓					
Lee et al (2010)(b) [39]		✓	✓		✓	✓					
Turkbey et al (2010) [40]	✓	✓		✓	✓	✓[c]	✓		✓		
Brown et al (2009) [56]	✓		✓		✓	✓					
Zhang et al (2009) [80]	✓	✓	✓		✓	✓					
Cirillo et al (2008) [57]			✓		✓	✓					

Abbreviations: BC, body coil; DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; ER, endorectal coil; PC, pelvic/phased-array coil; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopic imaging; T, Tesla.

[a] Staging protocol (n=89 patients) used ER plus pelvic coil, whereas detection protocol (n=44 patients) used pelvic coil only.

[b] Cardiac coil.

[c] Torso array coil.

Note 1. "✓" indicates presence of characteristic whereas a shaded cell indicates the absence of the column characteristic and "+" indicates the addition of.

Note 2. Experienced reader defined as any experience such as: "specializes" or "experience" or "dedicated" or "radiology specialist" for example, as stated in original paper.

Table 4. Outcome: Diagnostic Accuracy (N=49 Studies)

Study (Year)	% Sensitivity (95% CI)	% Specificity (95% CI)	% PPV (95% CI)	% NPV (95% CI)
Orczyk et al (2016) [42]	MRI+DCE+DWI: 52.6	MRI+DCE+DWI: 96.5	MRI+DCE+DWI: 87.2	MRI+DCE:DWI: 81.6
Bittencourt et al (2015)[a] [43]	MRI_{EPE}: 98.0 (96.0-100.0) $MRI+DCE_{EPE}$: 82.0 (68.0-90.0) $MRI+DWI_{EPE}$: 95.0 (86.0-99.0) MRI: 68.0 MRI+DCE: 55.0 MRI+DWI: 80.0	MRI_{EPE}: 32.0 (24.0-39.0) $MRI+DCE_{EPE}$: 33.0 (22.0-44.0) $MRI+DWI_{EPE}$: 27.0 (18.0-39.0) MRI: 68.0 MRI+DCE: 73.0 MRI+DWI: 55.0	MRI_{EPE}: 54.0 (46.0-63.0) $MRI+DCE_{EPE}$: 50.0 (39.0-60.0) $MRI+DWI_{EPE}$: 52.0 (42.0-61.0) MRI: 64.0 MRI+DCE: 62.0 MRI+DWI: 57.0	MRI_{EPE}: 96.0 (92.0-99.0) DCE_{EPE}: 69.0 (50.0-83.0) DWI_{EPE}: 87.0 (66.0-97.0) MRI: 72.0 MRI+DCE: 66.0 MRI+DWI: 76.0
Boesen et al (2015)[a] [44]	$MRI+DCE+DWI_{EPE}$: 81.0 (63.0-93.0) $MRI+DCE+DWI_{SVI}$: 80.0 (29.0-97.0)	$MRI+DCE+DWI_{EPE}$: 78.0 (66.0-88.0) $MRI+DCE+DWI_{SVI}$: 99.0 (93.0-99.0)	$MRI+DCE+DWI_{EPE}$: 68.0 (50.0-82.0) $MRI+DCE+DWI_{SVI}$: 80.0 (29.0-97.0)	$MRI+DCE+DWI_{EPE}$: 88.0 (76.0-95.0) $MRI+DCE+DWI_{SVI}$: 99.0 (93.0-99.0)
Counago et al (2015)[b] [45]	MRI+DCE+DWI: 70.0 (41.6-98.4)	MRI+DCE+DWI: 93.8 (88.5-99.1)	MRI+DCE+DWI: 58.3 (30.4-86.2)	MRI+DCE+DWI: 96.2 (91.9-100)
de Cobelli et al (2015) [46]	$MRI+DCE+DWI_{EPE}$: 100.0 (93.0-100.0) $MRI+DCE+DWI_{SVI}$: 100.0 (59.0-100.0)	$MRI+DCE+DWI_{EPE}$: 10.0 (6.0-15.0) $MRI+DCE+DWI_{SVI}$: 8.0 (4.0-12.0)	$MRI+DCE+DWI_{EPE}$: 25.0 (20.0-32.0) $MRI+DCE+DWI_{SVI}$: 3.0 (1.0-7.0)	$MRI+DCE+DWI_{EPE}$: 100.0 (79.0-100.0) $MRI+DCE+DWI_{SVI}$: 100.0 (79.0-100.0)
Feng et al (2015)[a] [47]	$MRI+DCE+DWI_{EPE}$: 84.6	$MRI+DCE+DWI_{EPE}$: 87.2	$MRI+DCE+DWI_{EPE}$: 66.7	$MRI+DCE+DWI_{EPE}$: 94.9
Ghafoori et al (2015) [4]	MRI_{SVI}: 97.0 (89.0-99.0)	MRI_{SVI}: 98.0 (94.0-99.0)	MRI_{SVI}: 94.0 (85.0-98.0)	MRI_{SVI}: 99.0 (96.0-99.0)
Razi et al (2015) [48]	MRI+MRS: 42.4	MRI+MRS: 93.6	MRI+MRS: 82.3	MRI+MRS: 69.8
Reisaeter et al (2015)[c] [49]	MRI_{PZ}: 0.50 (0.44-0.55) MRI_{TZ}: 0.13 (0.05-0.25) $MRI+DWI_{PZ}$: 0.56 (0.50-0.62) $MRI+DWI_{TZ}$: 0.17 (0.08-0.29) $MRI+DCE_{PZ}$: 0.51 (0.45-0.57) $MRI+DCE_{TZ}$: 0.13 (0.05-0.25) $MRI+DCE+DWI_{PZ}$: 0.60 (0.55-0.66) $MRI+DCE+DWI_{TZ}$: 0.15 (0.07-0.27)	MRI_{PZ}: 0.79 (0.76-0.82) MRI_{TZ}: 0.94 (0.92-0.96) $MRI+DWI_{PZ}$: 0.83 (0.80-0.85) $MRI+DWI_{TZ}$: 0.94 (0.91-0.96) $MRI+DCE_{PZ}$: 0.83 (0.81-0.86) $MRI+DCE_{TZ}$: 0.97 (0.95-0.98) $MRI+DCE+DWI_{PZ}$: 0.80 (0.77-0.83) $MRI+DCE+DWI_{TZ}$: 0.95 (0.93-0.97)	MRI_{PZ}: 0.47 (0.42-0.55) MRI_{TZ}: 0.19 (0.08-0.36) $MRI+DWI_{PZ}$: 0.55 (0.49-0.60) $MRI+DWI_{TZ}$: 0.22 (0.11-0.38) $MRI+DCE_{PZ}$: 0.54 (0.48-0.59) $MRI+DCE_{TZ}$: 0.32 (0.14-0.55) $MRI+DCE+DWI_{PZ}$: 0.53 (0.48-0.58) $MRI+DCE+DWI_{TZ}$: 0.24 (0.11-0.42)	MRI_{PZ}: 0.81 (0.78-0.83) MRI_{TZ}: 0.91 (0.89-0.94) $MRI+DWI_{PZ}$: 0.83 (0.81-0.86) $MRI+DWI_{TZ}$: 0.91 (0.89-0.94) $MRI+DCE_{PZ}$: 0.82 (0.79-0.84) $MRI+DCE_{TZ}$: 0.91 (0.89-0.94) $MRI+DCE+DWI_{PZ}$: 0.84 (0.81-0.87) $MRI+DCE+DWI_{TZ}$: 0.91 (0.89-0.94)
Rud et al (2015) [41]	MRI+DWI: 73.0 (63.0-81.0)	MRI+DWI: 65.0 (54.0-74.0)	NR	NR
Schieda et al (2015)[d] [3]	$MRI+DCE+DWI_{EPE}$: 59.5 (49.1-68.2)	$MRI+DCE+DWI_{EPE}$: 68.0 (50.5-82.6)	NR	NR
Woo et al (2015)[e] [50]	MRI_{EPE}: 30.0 (17.9-44.6) $MRI+DWI_{EPE}$: 92.0 (80.8-97.8)	MRI_{EPE}: 92.5 (83.4-97.5) $MRI+DWI_{EPE}$: 55.2 (42.6-67.4)	NR	NR
Counago et al (2014) [5]	$MRI+DCE+DWI_{EPE}$: 57.1 (25.0-84.1)	$MRI+DCE+DWI_{EPE}$: 95.0 (83.0-98.0)	$MRI+DCE+DWI_{EPE}$: 66.6 (29.0-100)	$MRI+DCE+DWI_{EPE}$: 92.6 (85.0-100)
Gupta et al (2014) [6]	$MRI+DCE+DWI_{OCD}$: 81.6 $MRI+DCE+DWI_{EPE}$: 77.8	$MRI+DCE+DWI_{OCD}$: 86.4 $MRI+DCE+DWI_{EPE}$: 83.4	$MRI+DCE+DWI_{OCD}$: 91.2 $MRI+DCE+DWI_{EPE}$: 66.7	$MRI+DCE+DWI_{OCD}$: 73.1 $MRI+DCE+DWI_{EPE}$: 89.7
Kan et al (2014) [7]	$MRI+DWI_{EPE}$: 5.9	$MRI+DWI_{EPE}$: 94.9	$MRI+DWI_{EPE}$: 33.3	$MRI+DWI_{EPE}$: 69.8

Kim et al (2014)(a)[e] [8]	MRI+DWI$_{EPE}$: 75.0	MRI+DWI$_{EPE}$: 70.0	NR	NR
Kim et al (2014)(b)[f] [9]	MRI: 32.3 (25.5-37.2) MRI+DCE: 27.2 (22.7-32.6) MRI+DWI: 43.5 (38.8-50.8)	MRI: 94.0 (92.0-96.1) MRI+DCE: 93.2 (91.6-94.9) MRI+DWI: 91.1 (89.3-94.0)	MRI: 75.0 (71.6-86.2) MRI+DCE: 71.7 (67.3-79.9) MRI+DWI: 77.1 (72.2-85.0)	MRI: 66.2 (61.2-72.0) MRI+DCE: 63.2 (58.6-69.5) MRI+DWI: 69.8 (65.8-74.5)
Kitamura et al (2014) [10]	MRI: 48.0 DWI: 44.0 MRS: 29.0 MRI+DWI: 56.0 MRI+MRS: 62.0 MRI+DWI+MRS: 68.0	MRI: 81.0 DWI: 93.0 MRS: 86.0 MRI+DWI: 79.0 MRI+MRS: 65.0 MRI+DWI+MRS: 63.0	MRI: 68.0 DWI: 84.0 MRS: 62.0 MRI+DWI: 69.0 MRI+MRS: 60.0 MRI+DWI+MRS: 61.0	MRI: 65.0 DWI: 66.0 MRS:58.0 MRI+DWI: 68.0 MRI+MRS: 67.0 MRI+DWI+MRS: 70.0
Lawrence et al (2014)[c] [11]	MRI$_{EPE}$: 22.0 (10.0-41.0) MRI+DWI$_{EPE}$: 44.0 (27.0-64.0)	MRI$_{EPE}$: 85.0 (73.0-92.0) MRI+DWI$_{EPE}$: 83.0 (71.0-91.0)	MRI$_{EPE}$: 43.0 (20.0-69.0) MRI+DWI$_{EPE}$: 57.0 (36.0-76.0)	MRI$_{EPE}$: 68.0 (56.0-78.0) MRI+DWI$_{EPE}$: 75.0 (61.0-85.0)
Lista et al (2014) [12]	MRI$_{EPE}$: 33.0 MRI+DCE+DWI$_{EPE}$: 58.0 MRI$_{SVI}$: 50.0 MRI+DCE+DWI$_{SVI}$: 75.0	MRI$_{EPE}$: 96.0 MRI+DCE+DWI$_{EPE}$: 98.0 MRI$_{SVI}$: 94.0 MRI+DCE+DWI$_{SVI}$: 96.0	MRI$_{EPE}$: 84.0 MRI+DCE+DWI$_{EPE}$: 95.0 MRI$_{SVI}$: 66.0 MRI+DCE+DWI$_{SVI}$: 80.0	MRI$_{EPE}$: 69.0 MRI+DCE+DWI$_{EPE}$: 75.0 MRI$_{SVI}$: 89.0 MRI+DCE+DWI$_{SVI}$: 94.0
Otto et al (2014)[g] [13]	MRI+DCE+DWI(+MRS)$_{EPE}$: 90.0 MRI+DCE+DWI(+MRS)$_{SVI}$: 80.0	MRI+DCE+DWI(+MRS)$_{EPE}$: 74.0 MRI+DCE+DWI(+MRS)$_{SVI}$: 96.0	NR -	NR -
Park et al (2014)(a) [14]	MRI+DCE+DWI$_{T3}$: 55.9	MRI+DCE+DWI$_{T3}$: 82.2	MRI+DCE+DWI$_{T3}$: 59.1	MRI+DCE+DWI$_{T3}$: 80.2
Roethke et al (2014) [15]	MRI$_{SVI}$: 48.6	MRI$_{SVI}$: 97.7	MRI$_{SVI}$: 68.0	MRI$_{SVI}$: 94.9
Song et al (2014) [16]	MRI+DWI$_{ANT-T}$: 65.1 MRI+DWI$_{POST-T}$: 72.2	MRI+DWI$_{ANT-T}$: 88.8 MRI+DWI$_{POST-T}$: 78.9	MRI+DWI$_{ANT-T}$: 76.3 MRI+DWI$_{POST-T}$: 86.0	MRI+DWI$_{ANT-T}$: 82.1 MRI+DWI$_{POST-T}$: 61.2
Styles et al (2014)[h] [17]	MRI: 53.2 DWI: 62.3	NR -	NR -	NR -
Armitage et al (2013) [18]	MRI$_{EPE}$: 41.0 (23.0-61.0)	MRI$_{EPE}$: 69.0 (42.0-87.0)	MRI$_{EPE}$: 69.0	MRI$_{EPE}$: 41.0
Cerantola et al (2013) [19]	MRI+DCE+DWI$_{EPE}$: 35.0 (19.0-55.0)	MRI+DCE+DWI$_{EPE}$: 90.0 (73.0-98.0)	MRI+DCE+DWI$_{EPE}$: 79.0 (49.0-95.0)	MRI+DCE+DWI$_{EPE}$: 57.0 (41.0-71.0)
Hegde et al (2013) [20]	MRI+DCE+DWI$_{EPE}$: 28.0 MRI+DCE+DWI$_{SVI}$: 50.0	MRI+DCE+DWI$_{EPE}$: 91.0 MRI+DCE+DWI$_{SVI}$: 99.0	MRI+DCE+DWI$_{EPE}$: 50.0 MRI+DCE+DWI$_{SVI}$: 83.0	MRI+DCE+DWI$_{EPE}$: 79.0 MRI+DCE+DWI$_{SVI}$: 96.0
Hole et al (2013) [21]	MRI+DWI$_{EPE}$: 56.3	MRI+DWI$_{EPE}$: 82.2	MRI+DWI$_{EPE}$: 85.4	MRI+DWI$_{EPE}$: 50.4
Isebaert et al (2013) [22]	MRI: 25.1 (21.0-29.7) DCE: 22.8 (19.3-26.7) DWI: 36.8 (31.5-42.4) MRI+DCE: 35.6 (31.4-40.1) MRI+DWI: 44.8 (39.4-50.3) DCE+DWI: 43.7 (38.6-48.9) MRI+DCE+DWI: 49.3 (44.2-54.4)	MRI: 94.7 (92.2-96.5) DCE: 94.2 (92.2-95.7) DWI: 93.8 (90.9-95.8) MRI+DCE: 90.3 (87.3-92.7) MRI+DWI: 90.2 (86.8-92.9) DCE+DWI: 89.0 (85.5-91.7) MRI+DCE+DWI: 86.5 (82.6-89.6)	MRI: 86.9 (80.2-91.6) DCE: 84.3 (78.7-88.6) DWI: 89.1 (83.5-93.0) MRI+DCE: 83.8 (77.9-88.3) MRI+DWI: 86.7 (80.9-90.9) DCE+DWI: 84.7 (79.1-89.0) MRI+DCE+DWI: 83.7 (77.9-88.2)	MRI: 49.9 (44.2-55.5) DCE: 49.0 (43.2-54.8) DWI: 53.5 (48.1-58.7) MRI+DCE: 52.0 (46.3-57.7) MRI+DWI: 55.6 (50.3-60.7) DCE+DWI: 55.0 (49.6-60.3) MRI+DCE+DWI: 56.6 (51.3-61.7)
Jeong et al (2013) [23]	MRI+DWI$_{EPE}$: 43.0 MRI+DWI$_{SVI}$: 34.9	MRI+DWI$_{EPE}$: 84.2 MRI+DWI$_{SVI}$: 93.8	MRI+DWI$_{EPE}$: 78.6 MRI+DWI$_{SVI}$: 62.4	MRI+DWI$_{EPE}$: 52.2 MRI+DWI$_{SVI}$: 83.1

	$MRI+DWI_{LNM}$: 14.0	$MRI+DWI_{LNM}$: 96.9	$MRI+DWI_{LNM}$: 22.9	$MRI+DWI_{LNM}$: 94.5
Johnston et al (2013) [24]	MRI_{EPE}: 20.0	MRI_{EPE}: 80.2	NR	NR
	MRI_{SVI}: 0	MRI_{SVI}: 94.2	-	-
Nepple et al (2013) [25]	MRI_{EPE}: 14.0	MRI_{EPE}: 88.0	MRI_{EPE}: 27.0	MRI_{EPE}: 76.0
	MRI_{SVI}: 38.0	MRI_{SVI}: 99.0	MRI_{SVI}: 75.0	MRI_{SVI}: 94.0
Pak et al (2013) [26]	$MRI+DWI_{EPE}$: 62.5	$DWI+MRI_{EPE}$: 82.1	$DWI+MRI_{EPE}$: 18.0	$DWI+MRI_{EPE}$: 97.2
Porcaro et al (2013) [27]	MRI_{EPE}: 76.0	MRI_{EPE}: 95.0	MRI_{EPE}: 91.0	MRI_{EPE}: 92.0
	MRI_{EPE}: 78.0	MRI_{EPE}: 96.0	MRI_{EPE}: 86.0	MRI_{EPE}: 92.0
	MRI_{SVI}: 88.0	MRI_{SVI}: 98.0	MRI_{SVI}: 82.0	MRI_{SVI}: 99.0
Pugh et al (2013) [28]	MRI_{EPE}: 50.9	MRI_{EPE}: 83.3	MRI_{EPE}: 60.4	MRI_{EPE}: 77.2
Renard-Penna et al (2013)[i] [30]	$MRI+DCE_{EPE}$: 81.3	$MRI+DCE_{EPE}$: 94.1	$MRI+DCE_{EPE}$: 72.2	$MRI+DCE_{EPE}$: 96.4
Roethke et al (2013) [29]	MRI_{T2}: 91.8	MRI_{T2}: 41.5	-	-
	MRI_{T3}: 40.7	MRI_{T3}: 92.9	-	-
	MRI_{T4}: 33.2	MRI_{T4}: 98.9	-	-
	MRI_{EPE}: 41.5	MRI_{EPE}: 91.8	MRI_{EPE}: 69.0	MRI_{EPE}: 78.0
Somford et al (2013) [31]	$MRI+DCE+DWI_{EPE}$: 58.2	$MRI+DCE+DWI_{EPE}$: 89.1	$MRI+DCE+DWI_{EPE}$: 84.1	$MRI+DCE+DWI_{EPE}$: 68.3
Soylu et al (2013)[i] [32]	MRI_{SVI}: 65 (45.5-84.6)	MRI_{SVI}: 90.7 (84.8-96.0)	MRI_{SVI}: 60 (40.0-79.3)	MRI_{SVI}: 92.5 (87.4-97.1)
	$MRI+DWI_{SVI}$: 78.0 (60.0-94.4)	$MRI+DWI_{SVI}$: 96.3 (92.3-99.1)	$MRI+DWI_{SVI}$: 82.0 (64.0-95.8)	$MRI+DWI_{SVI}$: 95.4 (91.2-99.1)
	$MRI+DCE+DWI_{SVI}$: 78.0 (60.0-94.4)	$MRI+DCE+DWI_{SVI}$: 96.3 (92.4-99.1)	$MRI+DCE+DWI_{SVI}$: 82.0 (64.7-95.8)	$MRI+DCE+DWI_{SVI}$: 95.4 (91.2-99.1)
Tanaka et al (2013)[j] [33]	$MRI+DCE+DWI_{EPE}$: 60.0	$MRI+DCE+DWI_{EPE}$: 86.0	$MRI+DCE+DWI_{EPE}$: 42.9	$MRI+DCE+DWI_{EPE}$: 92.5
Turkbey et al (2013)[k] [34]	$MRI+DCE+DWI+MRS_{EPE}$: 76.0	$MRI+DCE+DWI+MRS_{EPE}$: 95.0	$MRI+DCE+DWI+MRS_{EPE}$: 90.0	$MRI+DCE+DWI+MRS_{EPE}$: 88.0
	$MRI+DCE+DWI+MRS_{SVI}$: 67.0	$MRI+DCE+DWI+MRS_{SVI}$: 100	$MRI+DCE+DWI+MRS_{SVI}$: 100	$MRI+DCE+DWI+MRS_{SVI}$: 98.0
Tsao et al (2013) [35]	$MRI+DCE_{OCD}$: 63.1	$MRI+DCE_{OCD}$: 41.4	NR	NR
	$MRI+DCE_{EPE}$: 25.0	$MRI+DCE_{EPE}$: 70.7	-	-
	$MRI+DCE_{SVI}$: 35.3	$MRI+DCE_{SVI}$: 96.1	-	-
Kim et al (2012) [36]	ER-MRI_{EPE}: 33.3	ER-MRI_{EPE}: 96.6	ER-MRI_{EPE}: 91.7	ER-MRI_{EPE}: 56.9
	ER-MRI_{SVI}: 46.2	ER-MRI_{SVI}: 92.0	ER-MRI_{SVI}: 60.0	ER-MRI_{SVI}: 86.8
	PC-MRI_{EPE}: 31.3	PC-MRI_{EPE}: 97.5	PC-MRI_{EPE}: 93.8	PC-MRI_{EPE}: 54.2
	PC-MRI_{SVI}: 42.9	PC-MRI_{SVI}: 92.5	PC-MRI_{SVI}: 64.3	PC-MRI_{SVI}: 83.8
McClure et al (2012)[l] [51]	$MRI+DCE+DWI+MRS_{T3}$: 50.0	$MRI+DCE+DWI+MRS_{T3}$: 97.5	$MRI+DCE+DWI+MRS_{T3}$: 50.0	$MRI+DCE+DWI+MRS_{T3}$: 97.5
Novis et al (2011) [37]	$MRI_{T1\text{-}T2}$: 71.5	$MRI_{T1\text{-}T2}$: 58.9	$MRI_{T1\text{-}T2}$: 76.6	$MRI_{T1\text{-}T2}$: 52.4
	$DCE_{T1\text{-}T2}$: 67.2	$DCE_{T1\text{-}T2}$: 65.7	$DCE_{T1\text{-}T2}$: 79.3	$DCE_{T1\text{-}T2}$: 50.6
	$MRS_{T1\text{-}T2}$: 70.4	$MRS_{T1\text{-}T2}$: 58.7	$MRS_{T1\text{-}T2}$: 78.4	$MRS_{T1\text{-}T2}$: 48.2
	MRI_{EPE}: 50.0	MRI_{EPE}: 77.6	MRI_{EPE}: 13.7	MRI_{EPE}: 95.6
	MRI_{SVI}: 40.0	MRI_{SVI}: 83.1	MRI_{SVI}: 15.4	MRI_{SVI}: 94.7
Turkbey et al (2011) [38]	MRI: 58.0	MRI: 93.0	MRI: 70.0	MRI: 90.0
	DCE: 38.0	DCE: 98.0	DCE: 86.0	DCE: 87.0

	DWI: 53.0 MRS: 16.0	DWI: 95.0 MRS: 100	DWI: 73.0 MRS: 93.0	DWI: 89.0 MRS: 83.0
Lee et al (2010)(b) [39]	ER-MRI_{EPE}: 31.8 ER-MRI_{SVI}: 50.0 PC-MRI_{EPE}: 29.6 PC-MRI_{SVI}: 50.0	ER-MRI_{EPE}: 95.5 ER-MRI_{SVI}: 92.5 PC-MRI_{EPE}: 90.0 PC-MRI_{SVI}: 97.7	ER-MRI_{EPE}: 87.5 ER-MRI_{SVI}: 40.0 PC-MRI_{EPE}: 80.0 PC-MRI_{SVI}: 66.7	ER-MRI_{EPE}: 58.3 ER-MRI_{SVI}: 94.9 PC-MRI_{EPE}: 48.7 PC-MRI_{SVI}: 95.5
Turkbey et al (2010) [40]	MRI_{PZ+TZ}: 44.0 (37.0-51.0) DCE_{PZ+TZ}: 19.0 (14.0-25.0) MRS_{PZ+TZ}: 12.0 (7.0-18.0)	MRI_{PZ+TZ}: 83.0 (81.0-86.0) DCE_{PZ+TZ}: 95.0 (93.0-97.0) MRS_{PZ+TZ}: 97.0 (95.0-98.0)	NR - -	NR - -

Abbreviations: ANT-T, anterior index tumours; CI, confidence interval; DCE, dynamic-contrast enhanced imaging for tumor detection; DWI, diffusion-weighted imaging; EPE, extraprostatic extension; ER, endorectal coil; LNM, lymph node metastasis; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopic imaging; NPV, negative predictive value; NR, not reported; OCD, organ-confined disease; PC, pelvic coil; POST-T, posterior index tumors; PPV, positive predictive value; PZ, peripheral zone; SVI, seminal vesicle invasion; TZ, transitional zone.

[a] For a PI-RADS score of ≥4 (Bittencourt et al, 2015); for PI-RADS score of ≥4 to be more conservative (vs. score of 3) (Boesen et al, 2015); for a PI-RADS score of ≥3 (Feng et al, 2015).

[b] Based on 90 patients.

[c] Observer 2 with more than 10 years of experience (Reisaeter et al, 2015); more experienced of the two readers (Lawrence et al, 2014).

[d] For PI-RADS ≥3 and subset of 65 patients with standardized reporting (Schieda et a, 2015).

[e] Tumour ADC value using a cutoff of 1.09 (Kim et al, 2014a) and 0.89 (Woo et al, 2015) × $10^{-3}/mm^2$ for presence of EPE.

[f] Clinically significant cancer is likely to be present (score 4); values are for scores ≥4 (vs. score of 3), according to the European Society of Urogenital Radiology.

[g] For reader A. Kappa between 2 readers was 0.89 per patient analysis.

[h] Per tumor analysis (n=77 tumours).

[i] For the experienced reader.

[j] By prostate side or lobe (right side or left side).

[k] Among 133 patients, there were 50 patients (37.6%) who had MRI+DCE+MRS and 83 patients (62.4%) who had MRI+DCE+DWI+MRS.

[l] 208 lobes.

Table 5. Outcome: Diagnostic Accuracy (Qualitative Summary) (N=4 Studies)

Study (Year)	Details
Junker et al (2015) [81]	For 3 T MRI compared with histopathologic findings, the detection of prostate cancer based on receiver operator characteristics were: AUC: 0.90 (95% CI: 0.87-0.94) for MRI alone; AUC: 0.92 (95% CI: 0.88-0.95) for MRI+DWI; and AUC: 0.85 (95% CI: 0.80-0.90) for MRI+DCE.
Jung et al (2012) [77]	For 1.5 T MRI+ER compared with histopathologic findings in detecting EPE and SVI, the AUC ranged from 0.69 to 0.70 for three independent readers.
Hwii Ko et al (2011) [78]	The number and percentage of patients that were correctly categorized for EPE, SVI, and stage T3 tumours between 3 T MRI and pathology were 84 (69.4%), 102 (84.3%), and 89 (73.6%), respectively.
Zhang et al (2009) [80]	For 1.5 T MRI+ER compared with histopathologic findings in detecting stage T3a or higher, the AUC ranged from 0.71 to 0.75 for two independent readers. With the addition of MRS, the AUC ranged from 0.74 to 0.75 for the two independent readers.

Abbreviations: AUC, area under the curve; CI, confidence interval; DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; EPE, extraprostatic extension; ER, endorectal coil; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; SVI, seminal vesicle invasion; T, Tesla.

Outcome: Diagnostic Accuracy (no endorectal coil)

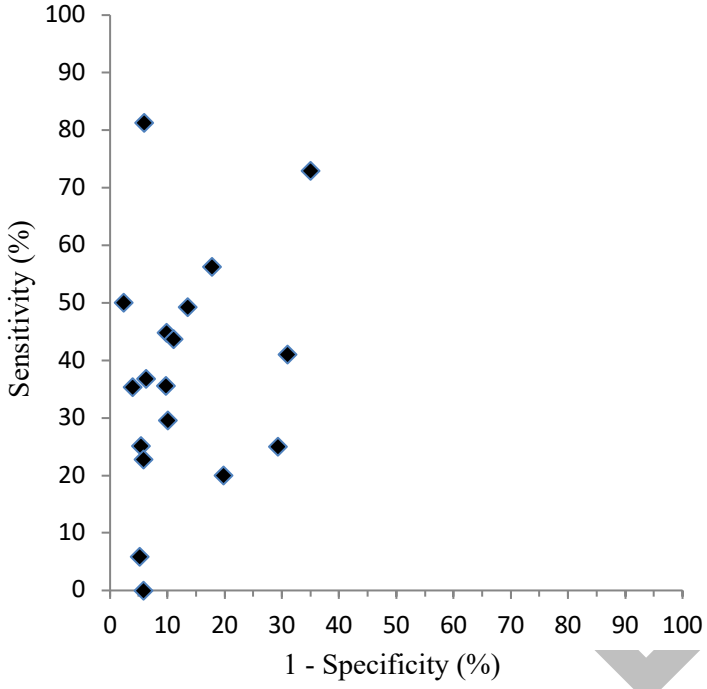


Figure 1. Diagnostic Accuracy of 1.5 T MRI ± DCE, DWI, MRS (n=18)

Abbreviations: DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; T, Tesla.

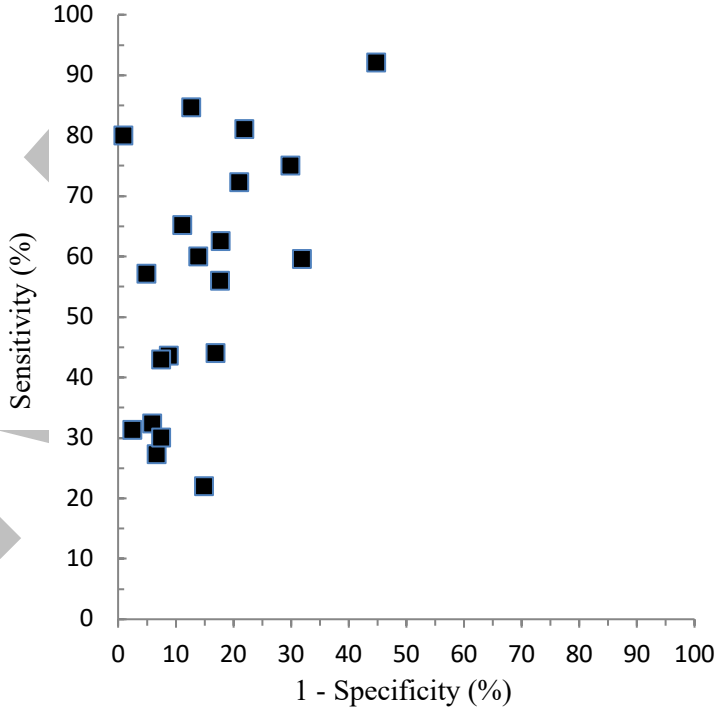


Figure 2. Diagnostic Accuracy of 3 T MRI ± DCE, DWI, MRS (n=14)

Abbreviations: DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; T, Tesla.

Table 6. Outcome: Change in TNM Stage Classification (N=20 Studies)

Study (Year)	SS	Clinical Stage [no. (%)][a]	MRI Stage [no. (%)][a]	Diff[b]	Results Summary[c]	Overall Interpretation of Findings
Counago et al (2015) [45]	274	T1a-T2a: 239 (87.2) T2b-c: 30 (11.0) T3-T4: 5 (1.8)	T1a-T2a: 67 (24.5) T2b-c: 157 (57.3) T3-T4: 50 (18.2)	-62.7 46.3 16.4	↓T1 ↑T2 ↑T3	Upstaging (validated in 58%-75%, see Appendix IV, Table 2)
Counago et al (2014) [5]	150	T1b-T2a: 127 (84.7) T2b-T2c: 9 (6.0) T3-T4: 0 (0)	T1b-T2a: 28 (18.7) T2b-T2c: 96 (64.0) T3-T4: 26 (17.3)	-66.0% 58.0% 17.3%	↓T1 ↑T2 ↑T3	Upstaging
Gupta et al (2014) [6]	60	T1: 52 (86.7) T2a/b: 7 (11.7)/T2c: 0 (0) [11.7] T3a: 0 (0)/T3b: 1 (1.6) [1.6]	T1: 0 (0) T2a/b: 11 (18.3)/T2c: 23 (38.3) [56.6] T3a: 21 (35)/T3b: 5 (8.4) [43.4]	-86.7% 44.9% 41.8%	↓T1 ↑T2 ↑T3	Upstaging
Lista et al (2014) [12]	85	T1c: 58 (68.2) T2a: 8 (9.4)/T2b: 3 (3.5)/T2c: 13 (15.3) [28.2] T3a: 3 (3.5)/T3b: 0 (0) [3.5]	T1c: 5 (5.9) T2a: 10 (11.8)/T2b: 8 (9.4)/T2c: 32 (37.6) [58.8] T3a: 16 (18.8)/T3b: 14 (16.5) [35.3]	-62.3% 30.6% 31.8%	↓T1 ↑T2 ↑T3	Upstaging
Park et al (2014)(b) [55]	282	T1c: 187 (66.3) T2a: 70 (24.8)/T2b: 16 (5.7)/T2c: 2 (0.7) T3a: 7 (2.5)/T3b: 0 (0) [2.5]	NR NR T3a: 96 (34.0)/T3b: 9 (3.2) [37.2]	NE NE 34.7%	- - ↑T3	Upstaging
Armitage et al (2013)[d] [18]	69	T1-T2a: 4 (5.8) T2b/c: 58 (84.1) T3-T4: 7 (10.1)	NR NR T3: 13 (37.1)	NE NE 27.0%	- - ↑T3	Upstaging
Cerantola et al (2013)[e] [19]	60	~T1: 24 (40.0) ~T2: 34 (56.7) ~T3: 2 (3.3) T4: 0 (0)	T1: 0 (0) T2a: 23 (38.3)/T2b: 11 (18.3)/T2c: 12 (20.0) [76.6] T3a: 12 (20.0) T4: 2 (3.3)	-40.0% 19.9% 16.7% 3.3%	↓T1 ↑T2 ↑T3 ↑T4	Upstaging (validated in 62%, see Appendix IV, Table 2)
Hegde et al (2013) [20]	118	T1c: 91 (77.1) T2a: 18 (15.3)/T2b: 6 (5.1)/T2c: 3 (2.5) [22.9] T3: 0 (0)	T1: 0 (0) T2: 102 (86.4) T3: 16 (13.6)	-77.1% 63.5% 13.6%	↓T1 ↑T2 ↑T3	Upstaging (validated, p = 0.0012, see Appendix IV, Table 2)

Hole et al (2013)[f] [21]	209	T1c: 96 (45.9) T2a/T2b: 53 (25.4)/T2c: 22 (10.5) [35.9] T3: 38 (18.2) T4: 0 (0)	T1: 0 (0) T2a/b: 52 (24.9)/T2c: 55 (26.3) [51.2] T3a: 74 (35.4)/T3b: 13 (6.2) [41.6] T4: 3 (1.4)	-45.9% 15.3% 23.4% 1.4%	↓T1 ↑T2 ↑T3 ↑T4	Upstaging (validated for 85%, see Appendix IV, Table 2)
Johnston et al (2013)[g] [24]	568	T1-T2a: 198 (34.9) T2b-T2c: 303 (53.3) T3-T45: 67 (11.8)	NR NR T3a: 113 (19.9)/T3b: 31 (5.5)	NE NE NE	- - -	-
Porcaro et al (2013) [27]	154	T1c: 42 (27.3) T2: 112 (72.7) T3: 0 (0)	T1: 0 (0) T2: 100 (64.9) T3a: 37 (24.0)/T3b: 17 (11.0) [35.0]	-27.3% -7.8% 35.0%	↓T1 ↓T2 ↑T3	Upstaging
Pugh et al (2013)[h] [28]	171	T1c: 87 (50.9) T2a: 43 (25.1)/T2b: 36 (21.1)/T2c: 5 (2.9) [49.1] T3: 0 (0)	- ≤T2: 123 (71.9) T3a-b: 48 (28.1)	-28.1% 22.8% 28.1%	↓T1 ↑T2 ↑T3	Upstaging
Renard-Penna et al (2013)[l] [30]	101	T1c: 75 (74.3) T2a: 8 (7.9)/T2b: 7 (6.9) [14.8] T3a: 6 (5.9)/ T3b: 5 (5.0) [10.9]	T1: 0 (0) T2: 83 (82.2) T3a: 13 (12.9)/T2b: 0 (0) [12.9]	-74.3% 67.4% 2.0%	↓T1 ↑T2 ↑T3	Upstaging
Somford et al (2013) [31]	183	T1: 95 (51.9) T2: 67 (36.6) T3: 21 (11.5)	T1: 0 (0) T2: 120 (65.6) T3: 63 (34.4)	-51.9% 29.0% 22.9%	↓T1 ↑T2 ↑T3	Upstaging
Tanaka et al (2013) [33]	67	T1: 0 (0) T2a: 28 (41.8)/T2b: 3 (4.5)/T2c: 20 (29.9) [76.2] T3a: 16 (23.9)	T1: 0 (0) T2: 53 (79.1) T3a: 14 (20.9)	0% 2.9% -3.0%	- ↑T2 ↓T3	Downstaging
Tsao et al (2013)[j] [35]	94	T1: 0 (0) ≤ T2: 94 (100) T3: 0 (0)	T1: 0 (0) T2: 58 (61.7) T3a: 27 (28.7)/T3b: 9 (9.6) [38.3]	0% -38.3% 38.3%	- ↓T2 ↑T3	Upstaging (validated in 11%-71%, see Appendix IV, Table 2)
Panebianco et al (2012)[k] [53]	105	T1c or T2a: 105 (100) T3: 0 (0)	T2: 73 (69.5) T3: 32 (30.5)	-30.5% 30.5%	↓T1c/2 ↑T3	Upstaging (validated in 28%, see Appendix IV, Table 2)

Ploussard et al (2011)[l] [54]	96	T1: ~84 (87.5)	T1 or T2: 68 (70.8)	-29.2%	↓T1/2	Upstaging
		T2: 12 (12.5)	-	NE	-	
		T3: 0 (0)	T3: 28 (29.2)	29.2%	↑T3	
Brown et al (2009)[m] [56]	62	T1c: 31 (54.4)	T1: 0 (0)	-54.4%	↓T1	Upstaging (validated in 38%-82%, see Appendix IV, Table 2)
		T2: 26 (50.0)	T2: 47 (75.8)	25.8%	↑T2	
		T3: 0 (0)	T3: 15 (24.2)	24.2%	↑T3	
Cirillo et al (2008) [57]	143	T1: 18 (12.6)	T1: 3 (2.1)	-10.5%	↓T1	Upstaging
		T2: 113 (79.0)	T2: 89 (62.2)	-16.8%	↓T2	
		T3: 12 (8.4)	T3: 49 (34.3)	25.9%	↑T3	
		T4: 0 (0)	T4: 2 (1.4)	1.4%	↑T4	

Abbreviations: DCE, dynamic contrast-enhanced imaging; diff, difference; DRE, digital rectal examination; DWI, diffusion-weighted imaging; EPE, extraprostatic extension; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; NE, not estimable; NR, not reported; SS, sample size.

[a] Staging according to the American Joint Committee on Cancer. Totals for a given stage are depicted by square brackets as percentages.

[b] Absolute within stage difference defined as: MRI stage minus clinical stage (baseline). A negative value indicates a decrease in the number or proportion of patients within a given stage category whereas a positive value indicates an increase in the number or proportion of patients within a given stage category. Change across stage categories is not specified.

[c] Reported for T stage overall.

[d] Subset of 35 patients had MRI results (Armitage et al, 2013).

[e] T1 given as "normal", T2 given as "palpable", and T3 given as "EPE" from DRE, as described in the original study.

[f] No tumor was identified at MRI for 10 patients and MRI was indeterminate for two patients.

[g] Patients categorized for clinical stage are part of the D'Amico risk classification reported in the original paper (e.g. low risk: T1-T2a, Gleason score ≤6, prostate-specific antigen level <10 ng/mL).

[h] T3a-b represents "EPE" from the original paper. Remainder of patients were assumed to be ≤T2 stage for MRI. Calculation of difference assumes: MRI ≤T2 stage minus clinical T2 stage (22.8%) and MRI ≤T2 stage minus clinical T2 stage minus clinical T1 stage (28.1%).

[i] Five patients unaccounted for in MRI results, therefore overall findings are uncertain.

[j] ≤T2 patients are categorized according to "clinically localized" as described in the study.

[k] Group A assigned to T2 and Group B assigned to T3 for MRI staging. Group A defined in the original paper as patients submitted to bilateral nerve-sparing radical prostatectomy. Group B defined in the original paper as patients submitted to unilateral nerve-sparing or non-nerve sparing radical prostatectomy.

[l] T1 or T2 in MRI subtracted from combined T1 and T2 in clinical stage (29.2%)

[m] Clinical stage only available for 57 patients.

Note 1. ~indicates that either stage or number was estimated or assumed from limited data in the original paper.

Table 7. Outcome: Change in Risk Stratification Category (Qualitative Summary) (N=2 Studies)

Study (Year)	Details
Counago et al (2015) [45]	Risk group changes were as follows: pre-MRI as 34.7% low risk, 46.4% intermediate risk, and 19.0% high risk. Post-MRI as 11.7% low risk, 59.5% intermediate risk, and 28.8% high risk. Overall, there were fewer individuals categorized as low risk and more individuals categorized as intermediate or high risk.
Counago et al (2014) [5]	Risk group change in 35 of 103 (33.9%) patients with a provisional treatment plan. There were 21 (20.1%) patients who were changed from low risk to intermediate risk, one (0.9%) patient that was changed from low risk to high risk, five (3.8%) patients that were changed from intermediate risk to intermediate-high risk, and eight (7.6%) patients that were changed from intermediate risk to high risk.

Abbreviations: DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging.

Table 8. Outcome: Change in Treatment Plan (N=6 Studies)

Study (Year)	MRI	Treatment Plan: Before MRI [no. (%)]	Treatment Plan: After MRI [no. (%)]	Diff[a]	Results	Overall	Correctness
Counago et al (2015) [45]	3 T MRI+DCE +DWI	High doses: 179 (65.3) ≥interm+high risk pts HT1: 92 (33.6) HT2: 125 (45.6)	High doses: 242 (88.3) ≥interm+high risk pts HT1: 116 (42.3) HT2: 171 (62.4)	23.0% 8.7% 16.8%	↑ therapy	More intense therapy	"less disagreement with therapy changes based on imaging compared to clinical staging therapy decision-making"
Rud et al (2015) [41]	1.5 T MRI+DWI	NR	Intervention arm: 59/222 pts had an altered procedure on one or both sides [cT1: 30 (51%), cT2: 29 (49%), rT3: 49 (83%)]	26.6%	Change	Authors stated: more radical excision for all surgical changes	NR
Counago et al (2014) [5]	3 T MRI+DCE +DWI	Low risk CTV: prostate; doses: 78 Gy; HT: none Low risk CTV: prostate; doses: 78 Gy; HT: none Interm risk CTV: prostate + SSVV; doses: 80 Gy; HT: none Interm risk CTV: prostate + SSVV; doses: 80 Gy; HT: none	Interm risk (21 pts) CTV: prostate + SSVV; doses: 80 Gy; HT: none High risk (1 pt) CTV: prostate + SSVV; doses: 80 Gy; HT x 24 mo Interm-high risk (5 pts) CTV: prostate + SSVV; doses: 80 Gy; HT x 6 mo High risk (8 pts) CTV: prostate + SSVV; doses: 80 Gy; HT x 24 mo	20.1% 0.9% 3.8% 7.6%	↑ risk group for all (35 pts)	More intense therapy	NR
Park et al (2014)(a) [14]	3 T MRI+DCE +DWI	NR	More aggressive surgery: 40/93 pts Less aggressive surgery: 53/93 pts	43.0% 57.0%	Change in both directions	More and less intense therapy	More aggressive surgery: 25/40 pts (63%) More conservative surgery: 48/53 pts (91%)
McClure et al (2012) [51]	1.5 T MRI+DCE +DWI+ MRS	NNS: 35/208 sides (16.8%) NS: 173/208 sides (83.2%)	Less aggressive surgery, change from NNS to NS: 17/28 pts (60.7%) More aggressive surgery, change from NS to NNS: 11/28 (39.3%)	27% pts had a change in surgery plan	Change in both directions	More and less intense therapy	Among initial NS surgical plan, 97.1% were T2 staged at imaging, of which 97.0% were pathologically confirmed as pT2.

Panebianco et al (2012) [53]	3 T MRI+DCE +DWI+ MRS	BL/NS/RP: 105 (100) UL/NS/RP: 0 (0) Non/NS/RP 0 (0)	BL/NS/RP: 73 (69.5) UL/NS/RP: 21 (20) Non/NS/RP: 11 (10.5)	-30.5% 20.0% 10.5%	↓ BL/NS/RP ↑ UL/NS/RP ↑ Non/NS/RP	More intense therapy	Histopathological confirmation of image-based treatment plans was appropriate in 70/73 (95.9%) of pts submitted for BL/NS/RP whereas the treatment plan to change to UL/NS/RP or Non/NS/RP after imaging (but before surgery) was appropriate in 28/32 (87.5%) pts.

Abbreviations: BL/NS, bilateral nerve-sparing surgery; CTV, clinical target volume; DCE, dynamic contrast-enhanced; diff, difference; DWI, diffusion-weighted imaging; Gy, gray; HT, hormone therapy; HT1, hormone therapy according to 'initial criteria'; HT2, hormone therapy according to MSKCC criteria; interm, intermediate; mo, months; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; MSKCC, Memorial Sloan-Kettering Cancer Centre; NNS, non-nerve sparing; Non/NS, non-nerve-sparing surgery; NR, not reported; NS, nerve-sparing; pts, patients; RP, radical prostatectomy; SSVV, seminal vesicles; T, Tesla; UL/NS, unilateral nerve-sparing surgery.

[a] Absolute difference within a treatment category defined as: After MRI treatment plan minus before MRI treatment plan. A negative value indicates a decrease in the number or proportion of patients within a given treatment category whereas a positive value indicates an increase in the number of proportion of patients within a given treatment category. Change across treatment categories is not specified.

Table 9. Outcome: Positive Surgical Margin Status (N=20 Studies)

Study (Year)	SS[a]	Type of RP Surgery: OS	LP	RB	PSA Level	Positive Surgical Margin Status [no. (%)]
Staging from imaging used to alter RP (N=7)						
Randomized Controlled Trial (N=1)						
Rud et al (2015) [41]	222/ 216	-	-	✓	Low	MRI: 43 (19.4) Non-MRI: 49 (22.7)
Observational Studies (N=6)						
Yao et al (2014) [75]	84	-	-	✓	Low	18 (21.4)
Nepple et al (2013)[b] [25]	91	✓	-	-	Low	29 (31.9)
Roethke et al (2013)[c] [29]	385	✓	-	-	Low	57 (14.8)
Tanaka et al (2013)[d] [33]	134	-	-	✓	Low	21 (15.7)
McClure et al (2012)[e] [51]	104	-	-	✓	Low	7 (6.7)
Panebianco et al (2012)[f] [53]	105	✓	-	-	Low	10 (9.5)
Overall Range:						6.7% - 31.9%
Staging from imaging *not* used to alter RP (N=13)						
Kitamura et al (2014) [10]	54	-	-	-	Low	23 (42.6)
Park et al (2014)(a) [14]	353	-	-	✓	Low	46 (13.0)
Song et al (2014)[g] [16]	382	✓	-	✓	Low	17 (4.5)
Cerantola et al (2013) [19]	60	✓	-	-	Intermediate	26 (43.3)
Hegde et al (2013) [20]	118	✓	-	✓	Low	30 (25.4)
Pak et al (2013)[h] [26]	944	✓	-	✓	Intermediate	140 (14.8)
Renard-Penna et al (2013)[i] [30]	101	✓	-	✓	Low	18 (17.8)
Guzzo et al (2012) [76]	170	-	-	-	Low	8 (4.7)
Novis et al (2011) [37]	35	✓	-	-	Low	11 (31.4)
Ploussard et al (2011)[j] [54]	96	-	✓	-	Low	13 (13.5)
Lee et al (2010)(a) [79]	67	✓[k]	-	-	Intermediate	21 (31.3)
Lee et al (2010)(b) [39]	91	✓	-	-	Intermediate	4 (4.4)
Brown et al (2009) [56]	62	✓	✓	-	Low	18 (29.0)
Overall Range:						4.4% - 43.3%

Abbreviations: DCE, dynamic contrast-enhanced; DWI, diffusion-weighted imaging; LP, laparoscopic surgery; MRI; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; OS, open surgery; PSA, prostate specific antigen; RB, robotic-assisted surgery; RP, radical prostatectomy; SS, study size.

[a] Intervention/control.

[b] Three cases aborted. Nerve-sparing surgery was performed at the single surgeon's discretion based on biopsy, PSA, and MRI results.

[c] Preoperative MRI was used in surgical planning (feasibility and extent of nerve-sparing RP).

[d] Per lobe analysis. MRI determined the feasibility and extent of nerve-sparing RP.

[e] Cases submitted for the type of nerve-sparing RP based on MRI results.

[f] The resection plan was reevaluated after review of the imaging results.

[g] Percentage different than what was reported in the original paper where the denominator was not clear.

[g] Per lobe analysis.

[i] 15 cases in pT2, 3 cases in pT3.

[j] Laparoscopic extraperitoneal RP.

[k] Radical perineal RP.

Note 1. Guzzo et al, 2012, type of RP technique used not reported.

Table 10. Outcome: Biochemical Recurrence (Qualitative Summary) (N=5 Studies)

Author (Year)	Details
Park et al (2014)(b) [55]	-After RP and a median follow-up of 26 mo, there were 61 of 282 patients (21.6%) that experienced BCR defined as an initial PSA ≥0.2 ng/mL and a second confirmatory PSA >0.2 ng/mL. -Multivariate Cox regression analysis showed 3 T MRI for tumour detection as a significant predictor of BCR after RP (HR: 2.38, 95% CI: 1.03 to 6.00, p=0.047) whereas EPE on 3 T MRI was not shown to be a significant predictor of BCR after RP (HR: 1.81, 95% CI, 0.93 to 3.54, p=0.081). For both analyses, what was used as the referent group was not clear.
Armitage et al (2013) [18]	-Of 35 patients who had a 1.5 T MRI and RP, nine (26%) had BCR at one year of follow-up after surgery. BCR was defined as PSA ≥0.4 ng/dL followed by another risk in level.
Jeong et al (2013) [23]	-Median BCR-free survival was 69 mo (5.8 yrs). The overall five-year BCR-free survival rate was 56.1%. Non-organ-confined disease on predominately 3 T MRI was independently associated with BCR (HR: 1.92, 95% CI: 1.49 to 2.49, p<0.001) in adjusted models. BCR was defined according to the guidelines of the American Urological Association Localized Prostate Cancer Update Panel report.
Ploussard et al (2011) [54]	-A 1.5 T MRI did not predict BCR recurrence-free survival, defined as a PSA level >0.2 ng/mL, after a mean follow-up time of 29 mo (log-rank test: T3 versus T1-T2, p=0.853).
Lee et al (2010)(a) [79]	-The proportion of patients that showed BCR (PSA ≥0.4 ng/mL) was higher among patients that were not downstaged (cT3 and ≥pT3a) from clinical staging to pathological staging compared with those patients that were downstaged (cT3 to ≤pT2c) using 3 T MRI (78.6% versus 17.0%, p<0.001). The other indices of BCR showed a similar statistically significant trend, favouring those patients that were downstaged compared with those patients that were not downstaged upon pathology review (median BCR-free interval (months), downstaged: 52.0 versus not downstaged: 8.5, p<0.001; three-year BCR-free probability (PSA <0.4 ng/mL), downstaged: 89% versus not downstaged: 36%, p<0.001).

Abbreviations: BCR, biochemical recurrence; CI, confidence interval; DCE, dynamic contrast-enhanced; DWI, diffusion-weighted imaging; EPE, extraprostatic extension; HR, hazard ratio; mo, months; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; PSA, prostate-specific antigen; RP, radical prostatectomy; yrs, years.

Appendix 8. Subanalysis

(1) Studies of Upstaging and Histopathological Confirmation

Q: Among studies that showed patients were upstaged by MRI ± DCE, DWI, MRS in Appendix 7, Table 6, additionally, what was the correctness of that upstaging found by histopathology, if performed?

Table 1. Correctness of Upstaging (N=7 Studies)

Study (Year)	Results
Counago et al (2015) [45]	Correct staging was shown in 64/90 (71.1%) pts. For 74.5% pts upstaged to T2b-c from cT1-T2a and for 58.3% pts upstaged to T3 from cT1-T2.
Cerantola et al (2014) [19]	Correct staging was confirmed in 37/60 (61.7%) pts. Among the 23/60 (38.3%) pts with incorrect staging, there were 3/60 (5%) pts that were over-staged to locally advanced tumours and 20/60 (33.3%) pts that were under-staged to organ-confirmed tumour.
Hedge et al (2013) [20]	Statistically significant differences were shown between MRI-staged T2 prostate cancers and T3 prostate cancers when compared to final pathology, with a trend towards more advanced tumours (pT3b) given by pathology to also have been observed as T3 on MRI compared to T2 (p = 0.0012), suggesting correct classification.
Hole et al (2013) [21]	Among 89 pts classified as locally advanced prostate cancer (T3/T4) by MRI, 13/89 (14.6%) pts were not confirmed as having EPE by histopathology. Whereas, EPE on histopathology was shown for 76/89 (85.4%) pts that were staged by MRI as pT3/pT4.
Tsao et al (2013) [35]	Of the T2 staged patients by MRI, 29.3% (17/58) were under-staged (pT3) and 70.7% (41/58) were correctly staged (pT2). Of the T3a staged patients by MRI, 77.8% (21/27) were over-staged (pT2), 11.1% (3/27) correctly staged as T3a (pT3a), and 11.1% under-staged (pT3b). Of the T3b staged patients by MRI, one-third (33.3%, 3/9) were over-staged (pT2) and two-thirds (66.7%, 6/9) were correctly staged (pT3b).
Panebianco et al (2012) [53]	A higher proportion of pT3a/b tumours were also identified as T3 on MRI (28.1%, Group B: unilateral nerve-sparing and non-nerve-sparing) compared to tumours classified as T2 on MRI (2.7%, Group A: bilateral nerve-sparing).
Brown et al (2009) [56]	Correct classification by MRI occurred 34/41 (82.9%) among pT2 prostate cancers and 8/21 (38.1%) among pT3 prostate cancers.

Abbreviations: DCE, dynamic contrast-enhanced imaging; DWI, diffusion-weighted imaging; MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; pts, patients.

(2) Quantitative Diagnostic Accuracy Studies

Q: Among studies with paired data on diagnostic accuracy shown in Figures 1 and 2, additionally, what were their definitions of positivity?

Table 2. Definitions of Tumor, EPE and SVI from Studies Using DCE, DWI, MRS[a] (N=10 Studies)

Author (Year)	Tumour	EPE	SVI
Gupta et al (2014) [6]	Suspected lesions were defined by characteristics used in previous studies and in clinical practice, including T2W hypointensity in the PZ relative to the normal PZ except in cases of possible Bp-related hemorrhage, **early or intense enhancement with contrast agent, early contrast agent washout, and restricted diffusion indicated by low ADC values.**	Irregular bulging of the capsule, obliteration of the rectoprostatic angle, and asymmetry or tumour involvement of the NVB. **mpMRI staging, T3a: high degree of suspicion for EPE (unilateral or bilateral) w/o involvement of adjacent structures. T3b: high degree of suspicion of SVI w/o involvement of adjacent structures.**	
Lista et al (2014) [12]	**T2W, DCE, DWI: Tumour in contact w/ capsule, tumour extension into periprostatic fat, obliteration of recto-bladder angle, asymmetry of NVB and SV infiltration.**		
Otto et al (2014) [13]		Five established criteria, T2W: (1) asymmetry of NVB (2) obliteration of the rectoprostatic angle (3) irregular bulging of the prostatic contour (4) low signal intensity in the rectoprostatic fat (5) overt extracapsular cancer. Infiltration of the prostate capsule was defined by a regularly delineated tumour contact with a length of at least 10 mm. If such a tumour contact showed irregular delineations or signal defects of the prostate capsule, w/o hypointense T2W areas in the periprostatic fat or NVB, this region was positive for EPE. Presence/absence of EPE was determined. Confidence level was rated on a 5-point scale. **DWI, DCE: used to rule out false-positive findings caused by hemorrhage or inflammation after Bp.**	At least one of the following: disruption or loss of normal vesicle architecture, focal or diffuse areas of low signal intensity within the vesicles, asymmetric thickening or irregular shape of the vesicle wall + evident tumour at the prostate base extending into the SV. Confidence level in the presence/absence was rated on a 5-point scale. **DWI, DCE: used to rule out false-positive findings caused by hemorrhage or inflammation after Bp.**
Cerantola et al (2013) [19]	Positive for tumour if at least 2/3 criteria met: (1) low signal intensity on T2W images (**2) homogeneous enhancing lesion (3) reduced diffusion coefficient.**	An irregular capsule bulge, a periprostatic fat infiltration, an obliteration of the retroprostatic angle and/or an asymmetry of NVB.	
Hegde et al (2013) [20]	Signal intensity on T1W, T2W imaging, **presence of restricted diffusion on DWI, presence of contrast enhancement on DCE** combined + qualitative images/anatomic location. Interpretation included standard assessment, conventional MRI morphologic signs, of focal tumour, presence of EPE, NVB invasion, and SVI (among other LN and metastatic features). Qualitative designation of possible, probable, or definite EPE ± NVB or SVI was considered T3 positive. Otherwise, considered a T2 lesion.		

Jeong et al (2013) [23]	The radiologist decided whether the patient had EPE, SVI (and/or LNM) based on their medical training and on previously described features on MRI images.		
Somford et al (2013) [31]	**Localization of prostate cancer using all sequences, however presence of EPE was determined on T2W.**	T2W: according to established criteria for EPE and based on personal training and knowledge. Presence/absence was determined.	
Soylu et al (2013) [32]			T2W: based on criteria previously reported in the literature. **T2W+DWI: at least one of the following, disruption or loss of normal architecture of the SV, focal or diffuse areas of low signal intensity within the SV (w/o corresponding high signal intensity on T1W images at the same location), asymmetric thickening or irregularity of the SV wall, and evident tumour at the prostate base extending to the SV. T2W+DWI+DCE: at least one of the following, focal areas of enhancement within the SV, asymmetric or irregular SV wall thickening or enhancement, and evident tumour at the base of the prostate extending to the SV.** A 5-point ordinal scale was used, with a score of ≥3 defining positivity.
Turkbey et al (2013) [34]		**An imaging score was assigned to each lesion on the basis of its features on images obtained with different pulse sequences at MR imaging. T2W+DWI: the criterion for a "visible" lesion was a well-circumscribed, round-ellipsoid, low signal intensity region within the prostate gland. MRS: choline/citrate ratios within voxels in the biopsy core sites, abnormal ratio if ≥3 SD higher than the mean healthy ratio. DCE: direct visual interpretation of raw DCE T1W, criterion being a focus of early and intense enhancement with rapid washout compared with the background.**	
McClure et al (2012) [51]	Presence of tumour on T2W: round, ovid, or irregular areas of low signal intensity.	Presence of low signal intensity in the PZ of the prostate with irregular bulging or bowing of the prostate capsule at T2W, disruption of low signal intensity periprostatic band at T2W imaging, or direct involvement of the NVB and/or obliteration of the retroprostatic angle. A composite score based on T2W, **DWI, DCE, MRS** was reported as low, intermediate, or high for the likelihood of EPE on each side. Lesions were considered suspicious when at least one of the following occurred: **restricted diffusion, early or intense enhancement, washout, or elevated choline-to-citrate ratio.**	

Abbreviations: ADC, apparent diffusion coefficient; Bp, biopsy; DCE, dynamic-contrast enhanced imaging; DWI, diffusion-weighted imaging; EPE, extraprostatic extension; LN(M), lymph nodes (metastasis); MRI, magnetic resonance imaging ± DCE, DWI, MRS; MRS, magnetic resonance spectroscopic imaging; NVB, neurovascular bundle; PZ, peripheral zone; SD, standard deviations; SV(I), seminal vesicle (invasion); T1W, T1-weighted images; T2W, T2-weighted images; w/, with; w/o, without.

[a] Bolded text highlights where definitions are explicitly linked to the DCE, DWI, MRS technology.

Note 1. Grey shaded boxes indicate no information.

Appendix 9. Quality of Evidence

Table 1. Quality Assessment by QUADAS-2[a] (N=19 Studies)

Study	Risk of Bias				Applicability Concerns		
	Patient Selection[b]	Index Test	Reference Standard[e]	Flow and Timing	Patient Selection	Index Test	Reference Standard
Ghafoori et al (2015) [4]	☹/☺	☹[c]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Gupta et al (2014) [6]	☹/☺	☹/☺[d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Lista et al (2014) [12]	☹/☺	☹[c,d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Otto et al (2014) [13]	☺	☹/☺[d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Roethke et al (2014) [29]	☹/☺	☹[c]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Cerantola et al (2013) [19]	☺	☹[c,d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Hegde et al (2013) [20]	☹/☺	☹[c,d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Jeong et al (2013) [23]	☺	☹[c,d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Nepple et al (2013) [25]	☺	☹[c]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Porcaro et al (2013) [27]	☹/☺	☺	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Pugh et al (2013) [28]	☺	☹[c]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Roethke et al (2013) [29]	☹/☺	☹[c]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Somford et al (2013) [31]	☺	☹/☺[d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Soylu et al (2013) [32]	☺	☹/☺[d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Turkbey et al (2013) [34]	☹/☺	☹/☺[d]	☺/ ?[g]	☺	☺	☺	☺
Kim et al (2012) [36]	☺	☹[c]	☺/ ?[g]	☺	☺	☺	☺
McClure et al (2012) [51]	☺	☹/☺[d]	☺/ ?[g]	☺	☺	☺	☺
Novis et al (2011) [37]	☹/☺	☹/☺[d]	☺/☹[f]/ ?[g]	☺	☺	☺	☺
Lee et al (2010)(b) [39]	☺	☺	☺/ ?[g]	☺	☺	☺	☺
Overall:	☹/☺	☹/☺	☺/☹/ ?	☺	☺	☺	☺

[a] ☺ low risk, ☹ high risk, ? unclear risk.
[b] High risk of bias for studies that did not report consecutively enrolling eligible patients or performing a random sample of eligible patients.
[c] High risk of bias for lack of blinding during interpretation of index test.
[d] High risk of bias due to the use of functional sequences without standardized practices and definitions in place for its use.
[e] All studies used histological confirmation by radical prostatectomy as the referent gold standard for determining tumor stage.
[f] Whether pathologists were blinded to index test results was not always known.
[g] Pathology criteria used to evaluate EPE and SVI was not known or not well-described in most studies.

Table 2. Supplementary Data Abstraction for QUADAS-2 (N=19 Studies)

Study (Year)	Time between imaging and RP	Blinding[a]	Pathology Evaluation
Ghafoori et al (2015) [4]	0-3 w	NR	The prostate cancer patients who did not have RP or with uncertain pathology data were excluded.
Gupta et al (2014) [6]	2 w	NR	Detailed pathological information was collected for the presence of EPE and SVI.
Lista et al (2014) [12]	NR	NR	Comprehensive assessment of prostate specimen and lymph nodes to determine the possible presence of EPE.
Otto et al (2014) [13]	0.3 w	NR	Final pathology report of both fresh frozen samples (intra-) and RP specimens (post-operatively).
Roethke et al (2014) [15]	NR	NR	Pathologic staging was reported according to 6th Ed UICC TNM staging system.
Cerantola et al (2013) [19]	NR	NR	At final pathology, presence of EPE, SVI were recorded.
Hedge et al (2013) [20]	7 w	NR	Evidence of EPE including SVI was recorded.
Jeong et al (2013) [23]	NR	NR	The presence of EPE and SVI were evaluated.
Nepple et al (2013) [25]	3 w	NR	NR
Porcaro et al (2013) [27]	NR	NR	EPE was defined as tumour extending into the prostate soft tissue. SVI was defined as tumour involvement of the muscular wall of the seminal vesicle.
Pugh et al (2013) [28]	NR	NR	EPE was defined as SVI or the presence of any malignant cell outside the prostatic capsule.
Roethke et al (2013) [29]	2 w	NR	The specimens were classified according to 2002 TNM staging classification.
Somford et al (2013) [31]	NR	NR	The presence of EPE was reported. Also the presence of SVI, defining a subset of the EPE cohort.
Soylu et al (2013) [32]	10 w	NR	SVI was defined microscopically as prostate cancer cell invasion into the wall of the seminal vesicle.
Turkbey et al (2013) [34]	7 w	Yes	EPE and SVI were assessed for each specimen.
Kim et al (2012) [36]	NR	Yes	Each cancer focus was determined by a genitourinary histopathologist.
McClure et al (2012) [51]	8 w	Yes	All reports were reviewed to determine presence of EPE or SVI.
Novis et al (2011) [37]	NR	NR	Tumour presence was recorded for each sextant, as well as the presence of EPE. SVI was recorded by side of involvement.
Lee et al (2010)(b) [39]	3 w	Yes	The presence of tumour cells beyond the capsular margin defined EPE.

Abbreviations: EPE, extraprostatic extension; NR, not reported; RP, radical prostatectomy; SVI, seminal vesicle invasion; TNM, tumour, node, metastasis; UICC, Union for International Cancer Control; w, weeks.
[a] Blinded to index test data.

Appendix 10. MpMRI and Technical Specifications

Use of T2-weighted magnetic resonance imaging (MRI) (± functional sequences) for staging can be performed at 1.5 T or 3 T, but 3 T is generally preferred to 1.5 T. Imaging parameters and spatial resolution should meet or exceed the minimum requirements defined in the current Prostate Imaging and Reporting and Data Systems Version 2 (PI-RADS v2) recommendations [1]. High-resolution, multi-planar T2-weighted MRI is the principal imaging sequence used for evaluation of extraprostatic extension (EPE). The functional sequences of diffusion-weighted imaging (DWI) and dynamic contrast-enhanced imaging (DCE), part of multiparametric MRI (mpMRI), are helpful in identifying the underlying tumour location. There is typically a desire to see the underlying tumour location concomitantly with the assessment of EPE. In addition, there is some limited evidence that the addition of functional sequences can improve the performance of local MRI staging in a recent meta-analysis (see Table 1 below) [52]. For this reason, mpMRI is suggested as the approach to local staging. The use of an endorectal coil remains controversial. The endorectal coil will produce images with higher signal-to-noise ratio; however, review of the published literature does not show a clear advantage when using an endorectal coil for staging (see Table 1 below) [52]. For this reason, endorectal coil is not recommended as an essential part of MRI for local staging. Caution should be exercised in relying on non-endorectal coil imaging at 1.5 T (current synthesis, Table 4-3 and below Table 1). There has been improvement in image quality with newer 1.5 T systems allowing for adequate diagnostic performance; however, on older 1.5 T systems, image quality may be insufficient. A quality assurance program will help ensure that whatever platform is used, good diagnostic performance similar to previously published results is achieved (approximately: sensitivity 60%, specificity 90%) (see below Table 1, All studies) [52]. This can be done by comparing results of pathological staging at prostatectomy to MRI staging.

Table 1. Diagnostic Accuracy By Study Characteristics*

	Sensitivity		Specificity	
Study Characteristics*	**Estimate**	**95% CI**	**Estimate**	**95% CI**
EPE				
All studies - overall	0.57	0.49-0.64	0.91	0.88-0.93
≥1 functional sequence	0.63	0.51-0.74	0.91	0.86-0.94
3 T + ER	0.60	0.40-0.78	0.88	0.82-0.92
3 T – ER	0.61	0.45-0.75	0.87	0.77-0.93
1.5 T + ER	0.55	0.45-0.65	0.93	0.89-0.95
1.5 T - ER	0.54	0.28-0.78	0.89	0.78-0.95
SVI				
All studies - overall	0.58	0.47-0.68	0.96	0.95-0.97
≥1 functional sequence	0.64	0.48-0.76	0.97	0.94-0.98
3 T + ER	0.45	0.30-0.60	0.97	0.92-0.99
3 T – ER	0.65	0.30-0.89	0.94	0.87-0.97
1.5 T + ER	0.62	0.51-0.71	0.97	0.95-0.98
1.5 T - ER	0.37	0.08-0.80	0.94	0.87-0.98

Abbreviations: CI, confidence interval; EPE, extraprostatic extension; ER, endorectal coil; SVI, seminal vesicle invasion; T, Tesla.
*Taken from De Rooij et al (2015) [52].