

*Division of Surgery and Interventional Science, University College London, London, UK; †Department of Urology, University College London Hospitals NHS Foundation Trust, London, UK; ‡Department of Urology, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; §Urology Service, Sidney Kimmel Center for Prostate and Urologic Cancers, Memorial Sloan-Kettering Cancer Center, New York, NY, USA; †Department of Urology, University Hospital Heidelberg, Heidelberg, Germany; ‡Department of Urology, University of California, Los Angeles, Los Angeles, CA, USA; §Department of Urology, Pitie-Salpêtrière Academic Hospital, Pierre et Marie Curie University, Paris, France; †Institut des Systèmes Intelligents et de Robotique, l'Université Pierre et Marie Curie, Paris, France; ‡Arthur Smith Institute for Urology and Departments of Radiology and Interventional Radiology, Hofstra North Shore-Jewish School of Medicine, New Hyde Park, NY, USA

Article info

Article history:
Accepted October 16, 2014

Keywords:
Image-guided biopsy
Image processing
Computer assisted
Magnetic resonance imaging
Prostate neoplasms
Software
Targeted biopsy

Abstract

Context: The current standard for diagnosing prostate cancer in men at risk relies on a transrectal ultrasound–guided biopsy test that is blind to the location of the cancer. To increase the accuracy of this diagnostic pathway, a software-based magnetic resonance imaging–ultrasound (MRI-US) fusion targeted biopsy approach has been proposed.

Objective: Our main objective was to compare the detection rate of clinically significant prostate cancer with software-based MRI-US fusion targeted biopsy against standard biopsy. The two strategies were also compared in terms of detection of all cancers, sampling utility and efficiency, and rate of serious adverse events. The outcomes of different targeted approaches were also compared.

Evidence acquisition: We performed a systematic review of PubMed/Medline, Embase (via Ovid), and Cochrane Review databases in December 2013 following the Preferred Reporting Items for Systematic reviews and Meta-analysis statement. The risk of bias was evaluated using the Quality Assessment of Diagnostic Accuracy Studies-2 tool.

Evidence synthesis: Fourteen papers reporting the outcomes of 15 studies (n = 2293; range: 13–582) were included. We found that MRI-US fusion targeted biopsies detect more clinically significant cancers (median: 33.3% vs 23.6%; range: 13.2–50% vs 4.8–52%) using fewer cores (median: 9.2 vs 37.1) compared with standard biopsy, respectively. Some studies showed a lower detection rate of all cancer (median: 50.5% vs 43.4%; range: 23.7–82.1% vs 14.3–59%). MRI-US fusion targeted biopsy was able to detect some clinically significant cancers that would have been missed by using only standard biopsy (median: 9.1%; range: 5–16.2%). It was not possible to determine which of the two biopsy approaches led most to serious adverse events because standard and targeted biopsies were performed in the same session. Software-based MRI-US fusion targeted biopsy detected more clinically significant disease than visual targeted biopsy in the only study reporting on this outcome (20.3% vs 15.1%).

Contributed equally.

Corresponding author. Division of Surgery and Interventional Science, University College London, London, UK, W1P 7NN. Tel. +44 20 3447 9194; Fax: +44 20 3447 9303.
E-mail address: massimo.valerio.12@ucl.ac.uk (M. Valerio).

http://dx.doi.org/10.1016/j.eururo.2014.10.026
0302-2838/© 2014 European Association of Urology. Published by Elsevier B.V. All rights reserved.
1. Introduction

Recent level 1 evidence has shown that men with intermediate- and high-risk prostate cancer (PCa) benefit from immediate radical therapy [1,2]. Therefore accurate attribution of cancer risk is critical. Inaccurate risk attribution can lead to both overtreatment and its consequent detrimental impact on quality of life for men with true low-risk cancer as well as undertreatment and the potential for missing the window of curability for those men with clinically significant disease.

The current standard for diagnosing PCa in men at risk relies on a transrectal ultrasound (TRUS)-guided biopsy test that is blind to the location of cancer. The test uses a random deployment of 10–12 needles to sample the prostate. If cancer is detected, a correct risk attribution is possible in approximately 50% [3,4]. To increase the accuracy of this diagnostic pathway, some experts advocate an increase in the number of cores using the same transrectal approach (saturation biopsies) [5]; others advocate transperineal template mapping biopsy (sampling the prostate every 5 mm) [6]. More recently, it was proposed that targeted biopsies to suspicious lesion(s) detected by multiparametric magnetic resonance imaging (mpMRI) may increase the diagnostic accuracy of TRUS biopsy [7]. A systematic review of the literature recently showed that image-guided targeted biopsy may detect the same number of men with clinically significant disease using fewer cores compared with standard biopsy [8].

The mpMRI involves combining T2-weighted images with diffusion-weighted images (DWIs) and dynamic contrast enhancement (DCE) [9,10]. A number of studies showed that mpMRI has high sensitivity and specificity [11–14]. Because the mpMRI and the biopsy are performed on different days with the latter commonly carried out using a transrectal ultrasound probe, a number of devices that use image-fusion software have been developed to overlay the mpMRI suspicious area onto the ultrasound (US) images at the time of biopsy.

We conducted a systematic review of published studies using MRI-TRUS image fusion targeted prostate biopsies to assess the accuracy of detection for clinically significant PCa compared with standard biopsies.

Conclusions: Software-based MRI-US fusion targeted biopsy seems to detect more clinically significant cancers deploying fewer cores than standard biopsy. Because there was significant study heterogeneity in patient inclusion, definition of significant cancer, and the protocol used to conduct the standard biopsy, these findings need to be confirmed by further large multicentre validating studies.

Patient summary: We compared the ability of standard biopsy to diagnose prostate cancer against a novel approach using software to overlay the images from magnetic resonance imaging and ultrasound to guide biopsies towards the suspicious areas of the prostate. We found consistent findings showing the superiority of this novel targeted approach, although further high-quality evidence is needed to change current practice.

© 2014 European Association of Urology. Published by Elsevier B.V. All rights reserved.
The primary outcome was the detection rate of clinically significant disease by MRI-TRUS image fusion targeted biopsy compared with the standard biopsy technique. The definition used to determine clinical significance was the one used by individual studies. Secondary outcomes were detection rate of all cancer, sampling efficiency and utility, and serious adverse event rate. We calculated sampling efficiency as the median number of cores to diagnose one man with clinically significant cancer. Utility was defined as the number of men with clinically significant disease detected by one sampling strategy but missed by the other strategy.

The data extraction form was designed according to recent Standards of Reporting for MRI-targeted Biopsy Studies (START) of the prostate [17]. The following information was extracted from each study: design, population (sample size, prior biopsy or treatment, age, prostate-specific antigen, prostate volume, number of mpMRI visible lesions per patient), mpMRI characteristics and interpretation, type of anaesthesia, standard biopsy (number of cores, sampling route, blinding), MRI-TRUS image fusion targeted biopsy procedure (software used, sampling route, time flow, number of cores per lesion), separate histologic outcomes for standard biopsy against targeted (detection rate of clinically significant disease, detection rate of all cancer, biopsy efficiency, utility of one biopsy approach compared with the other), and serious adverse events (classification used, number, and type). When the studies included a comparison between two targeted approaches, of which at least one was software based, the histologic outcomes mentioned earlier were also extracted for the alternative targeted strategy, and the direct comparison was displayed in a separate table.

2.2. Risk of bias in individual studies
Risk of bias in each study was assessed independently by the two investigators using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [18]. In case of disagreement, the senior author (H.U.A.) was consulted and arbitrated. The QUADAS-2 tool includes four domains—patient selection, index test, reference test, and time flow—that are all assessed in terms of risk of bias, and the first three in terms of applicability. The signalling questions used to score each domain were derived from the QUADAS-2 tool statement and are displayed in Supplementary Table 1.

2.3. Statistical analysis
Continuous variables are given using median, interquartile range (IQR), and overall range; the mean with standard deviation was used when the former was not available. Categorical variables are given using frequencies and percentages. All analyses were performed using SPSS v.20.0 (IBM Corp., Armonk, NY, USA).

3. Evidence synthesis
3.1. Risk of bias
Fourteen original papers were included in the final analysis (Fig. 1; Tables 1–3) [19–32]. Seven additional studies were presented only at congresses and are reported for completeness in Supplementary Table 2 [33–39].

Most studies had a low risk of bias and low applicability concerns with respect to patient selection (Figs. 2 and 3).
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Sample size</th>
<th>Prior biopsy</th>
<th>Age, yr</th>
<th>PSA, ng/mL</th>
<th>Prostate volume, mL</th>
<th>Strength of magnetic field, T</th>
<th>Sequences</th>
<th>Embryonal cell</th>
<th>Interpretation for consensus</th>
<th>Score used</th>
<th>Threshold for targeted</th>
<th>Time from MR to biopsy, d</th>
<th>Anaesthesia</th>
<th>Complications</th>
<th>MR-TRUS fusion targeted biopsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozer et al [10]</td>
<td>Paired cohort study</td>
<td>152</td>
<td>Biopsy naïve</td>
<td>Median [QR]: 63.7 (50.3–67.5)</td>
<td>Median [QR]: 6 (4–10)</td>
<td>Median [QR]: 38.5 (30–55)</td>
<td>1.5</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>Likert ≥2</td>
<td>30 (1.5–51); median: IQR 4</td>
<td>Local</td>
<td>TRUS 12 cores</td>
<td>Y</td>
<td>Urethane</td>
</tr>
<tr>
<td>De Jonge-Panis et al [21]</td>
<td>Paired cohort study</td>
<td>131</td>
<td>Biopsy naïve</td>
<td>Mean ± SD: 64.6 ± 6.7</td>
<td>Mean ± SD: 6.3 ± 4.1</td>
<td>Mean ± SD: 35.7 ± 35.1</td>
<td>1.5</td>
<td>T2W, DWI, DCE</td>
<td>Y</td>
<td>Y</td>
<td>Benign, intermediate, malignant</td>
<td>Intermediate NR</td>
<td>Local</td>
<td>TRUS 10–12 cores</td>
<td>NR</td>
<td>Virtual Navigation</td>
</tr>
<tr>
<td>De Jonge-Panis et al [21]</td>
<td>Paired cohort study</td>
<td>133</td>
<td>Biopsy naïve</td>
<td>Mean ± SD: 64.5 ± 7.9</td>
<td>Mean ± SD: 9 ± 3.9</td>
<td>Mean ± SD: 50 ± 20.6</td>
<td>1.5</td>
<td>T2W, DWI, DCE</td>
<td>Y</td>
<td>Y</td>
<td>Benign, intermediate, malignant</td>
<td>Intermediate NR</td>
<td>Local</td>
<td>TRUS 10–12 cores</td>
<td>NR</td>
<td>Urethane</td>
</tr>
<tr>
<td>Sont et al [21]</td>
<td>Paired cohort study</td>
<td>105</td>
<td>Previous negative biopsy</td>
<td>Median [QR]: 65 (59–70)</td>
<td>Median [QR]: 75 (5–11.2)</td>
<td>Median [QR]: 58 (36–82)</td>
<td>3</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>5-point scale ≥2</td>
<td>7–21</td>
<td>Local</td>
<td>TRUS 12 Antennas guided cores</td>
<td>N</td>
<td>Aronitis</td>
</tr>
<tr>
<td>Portalez et al [27]</td>
<td>Paired cohort study</td>
<td>129</td>
<td>Previous negative biopsy</td>
<td>Mean [range]: 64.7 (47–70)</td>
<td>Mean [range]: 64 (4–44)</td>
<td>Mean [range]: 75 (18–141)</td>
<td>1.5</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>NR</td>
<td>51 (31–57)</td>
<td>Local</td>
<td>TRUS 12 cores</td>
<td>N</td>
<td>Urethane</td>
</tr>
<tr>
<td>Miyagawa et al [31]</td>
<td>Paired cohort study</td>
<td>85</td>
<td>Previous negative biopsy</td>
<td>Median [QR]: 65 (56–84)</td>
<td>Median [QR]: 60 (3–4)</td>
<td>Median [QR]: 64.5 (46–95)</td>
<td>1.5</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>NR</td>
<td>N</td>
<td>Spinal</td>
<td>Local</td>
<td>TRUS virtual and transperineal combined for 10–11 cores</td>
<td>NR</td>
</tr>
<tr>
<td>Paech et al [24]</td>
<td>Paired cohort study</td>
<td>95</td>
<td>Mixed (54% biopsy naïve; 32% previous negative biopsy)</td>
<td>Mean [range]: 65 (46–78)</td>
<td>Mean [range]: 10.05 ± 1.8</td>
<td>Mean [range]: 56 ± 2.4</td>
<td>1.5</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>Likert ≥3</td>
<td>74 (4.9 ± 7.2)</td>
<td>Local</td>
<td>TRUS 12 cores</td>
<td>Y</td>
<td>Virtual Navigation</td>
</tr>
<tr>
<td>Weynck et al [35]</td>
<td>Paired cohort study</td>
<td>125</td>
<td>Mixed (54% biopsy naïve; 25% previous negative biopsy; 21% on active surveillance)</td>
<td>Mean [range]: 65 (35–71)</td>
<td>Mean [range]: 5.1 ± 3.7</td>
<td>Mean [range]: 46 (16–62)</td>
<td>3</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>5-point scale ≥2</td>
<td>NR</td>
<td>Local</td>
<td>TRUS 12 cores</td>
<td>Y</td>
<td>Aronitis</td>
</tr>
<tr>
<td>Mousavian et al [27]</td>
<td>Paired cohort study</td>
<td>13</td>
<td>Mixed (46% biopsy naïve; 31% previous negative biopsy; 23% on active surveillance)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>3</td>
<td>T2W, DWI, DCE with and without spectroscopy</td>
<td>Y</td>
<td>N</td>
<td>4-point scale</td>
<td>NR</td>
<td>2–3 wk (average with no other procedure)</td>
<td>NR</td>
<td>TRUS 10–12 cores</td>
<td>VR</td>
</tr>
<tr>
<td>Rand et al [28]</td>
<td>Paired cohort study</td>
<td>30</td>
<td>Mixed (43% biopsy naïve; 57% previous negative biopsy; 23% on active surveillance)</td>
<td>Mean [range]: 65 (40–87)</td>
<td>Mean [range]: 6.3 (5–8.8)</td>
<td>Mean [range]: 41 (11–50)</td>
<td>3</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>PI-RADS None</td>
<td>Median (range): 34 (8–30)</td>
<td>General or local</td>
<td>TRUS 12 cores</td>
<td>N</td>
<td>Urethane</td>
</tr>
<tr>
<td>Bommarito et al [34]</td>
<td>Paired cohort study</td>
<td>105</td>
<td>Mixed (33% biopsy naïve; 37% previous negative biopsy; 30% on active surveillance)</td>
<td>Mean [range]: 65.8 (42–87)</td>
<td>Mean [range]: 6.3 (6.5–8.6)</td>
<td>Mean [range]: 41 (10–50)</td>
<td>3</td>
<td>T2W, DWI, DCE</td>
<td>Y</td>
<td>Y</td>
<td>None</td>
<td>None</td>
<td>NR</td>
<td>Local</td>
<td>TRUS 12 cores</td>
<td>Y</td>
</tr>
<tr>
<td>Sehgal et al [34]</td>
<td>Paired cohort study</td>
<td>56</td>
<td>Mixed (59% previous negative biopsy; 41% on active surveillance)</td>
<td>Mean [range]: 63 ± 8.4</td>
<td>Mean [range]: 50 ± 11.3</td>
<td>Mean [range]: 50 ± 11.2</td>
<td>3</td>
<td>T2W, DWI, DCE, spectroscopy</td>
<td>Y</td>
<td>Y</td>
<td>Low, moderate, high</td>
<td>Low, moderate, high</td>
<td>None</td>
<td>Median: 39</td>
<td>Local</td>
<td>TRUS 12 cores</td>
</tr>
<tr>
<td>Sont et al [31]</td>
<td>Paired cohort study</td>
<td>171</td>
<td>Mixed (10% previous negative biopsy; 62% on active surveillance)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>3</td>
<td>T2W, DWI, DCE</td>
<td>N</td>
<td>N</td>
<td>5-point scale ≥2</td>
<td>7–21</td>
<td>Local</td>
<td>TRUS 12 Antennas guided cores</td>
<td>N</td>
<td>Aronitis</td>
</tr>
<tr>
<td>Ruff et al [32]</td>
<td>Comparative series</td>
<td>60</td>
<td>Mixed (32% biopsy naïve; 68% refugence; 101 failure after EBRT)</td>
<td>Median [range]: 64 (30–86)</td>
<td>Median [range]: 7 (1–74)</td>
<td>Median [range]: 24 (8–115)</td>
<td>1.5</td>
<td>T2W, DWI</td>
<td>N</td>
<td>N</td>
<td>Low, moderate, high</td>
<td>None</td>
<td>NR</td>
<td>Local</td>
<td>TRUS 12 cores</td>
<td>N</td>
</tr>
</tbody>
</table>

DCE = dynamic contrast enhanced; DWI = diffusion-weighted image; EBRT = external-beam radiation therapy; IQR = interquartile range; MR = magnetic resonance; N = no; NR = not reported; PSA = prostate-specific antigen; SD = standard deviation; T2W = T2 weighted; TRUS = transrectal ultrasound; Y = yes.

* Part of this cohort was already reported in another study included in this systematic review.

† The studies are ordered according to prior biopsy status with biopsy-naive population first.
Table 2 – Primary and secondary outcomes of magnetic resonance-transrectal ultrasound fusion targeted biopsy versus standard biopsy

<table>
<thead>
<tr>
<th>Study</th>
<th>Definition for clinically significant disease</th>
<th>Detection rate clinically significant disease SB, %</th>
<th>Detection rate clinically significant disease TB, %</th>
<th>Detection rate any cancer SB, %</th>
<th>Detection rate any cancer TB, %</th>
<th>Efficiency of SB in detecting one patient with clinically significant disease</th>
<th>Efficiency of TB in detecting one patient with clinically significant disease</th>
<th>Additional utility of SB vs TB, %</th>
<th>Additional utility of TB vs SB, %</th>
<th>Standard classification used</th>
<th>No. of serious adverse events SB, %</th>
<th>No. of serious adverse events TB, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozer et al [19]</td>
<td>Cancer core length ≥4 mm or Gleason ≥3+4</td>
<td>36.9</td>
<td>43.4</td>
<td>56.6</td>
<td>53.9</td>
<td>32.6</td>
<td>4.6</td>
<td>2.7</td>
<td>9.2</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Delongchamps et al [20], first study</td>
<td>Cancer core length ≥5 mm or Gleason ≥3+4</td>
<td>NR</td>
<td>NR</td>
<td>45.8</td>
<td>82.1</td>
<td>NR</td>
<td>NR</td>
<td>1.5</td>
<td>9</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Delongchamps et al [20], second study</td>
<td>Cancer core length ≥5 mm or Gleason ≥3+4</td>
<td>NR</td>
<td>NR</td>
<td>33.1</td>
<td>75.6</td>
<td>NR</td>
<td>NR</td>
<td>1.5</td>
<td>9.8</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Sonn et al [21]</td>
<td>Cancer core length ≥4 mm or Gleason ≥3+4</td>
<td>14.7</td>
<td>21.7</td>
<td>27.5</td>
<td>23.7</td>
<td>81.6</td>
<td>25.2</td>
<td>4.9</td>
<td>11.3</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Portalez et al [22]</td>
<td>Gleason ≥3+4</td>
<td>NR</td>
<td>NR</td>
<td>20.9</td>
<td>43.4</td>
<td>33.1</td>
<td>4.5</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Miyagawa et al [23]</td>
<td></td>
<td>NR</td>
<td>NR</td>
<td>40</td>
<td>52.9</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Puech et al [24]</td>
<td>Cancer core length ≥3 mm or Gleason ≥3+4</td>
<td>52</td>
<td>NR*</td>
<td>59</td>
<td>NR*</td>
<td>23.2</td>
<td>NR*</td>
<td>NR*</td>
<td>NR*</td>
<td>NR*</td>
<td>NR*</td>
<td>NR*</td>
</tr>
<tr>
<td>Wysock et al [25]</td>
<td>Gleason ≥3+4</td>
<td>NR</td>
<td>23.2</td>
<td>NR</td>
<td>36</td>
<td>31.2</td>
<td>9.8</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Kuru et al [26]</td>
<td>NCCN criteria</td>
<td>38</td>
<td>41.1</td>
<td>50.4</td>
<td>50.6</td>
<td>47.9</td>
<td>12.3</td>
<td>12.4</td>
<td>5.9</td>
<td>NR</td>
<td>Complications were given as a whole, with the two procedures performed in the same session: one surgical drainage of perineal haematoma; three interventions for haematuria</td>
<td></td>
</tr>
</tbody>
</table>

NR = not reported; SB = standard biopsy; TB = targeted biopsy.

* In this study, when the TB approach had already sampled one area, this area was excluded by the SB approach. In addition, the local histopathology analysis process did not allow us to distinguish discordance in grade between the two biopsy approaches (pooled analysis).

# In this study, the targeted approach, including visual and magnetic resonance-transrectal ultrasound fusion biopsy, is given as a whole; therefore it was not possible to distinguish the outcomes of software biopsy alone.
### Table 3 – Head-to-head comparison between two targeted biopsy approaches, one using an magnetic resonance–transrectal ultrasound fusion platform and the other a cognitive fusion approach

<table>
<thead>
<tr>
<th>Study</th>
<th>Targeted strategies used</th>
<th>Statistical analysis of the study</th>
<th>Histologic results</th>
<th>Detection rate</th>
<th>Detection rate</th>
<th>Detection rate</th>
<th>Efficiency of strategy 1</th>
<th>Efficiency of strategy 2</th>
<th>Additional strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poech et al [24]</td>
<td>1: Virtual Navigator</td>
<td>First cognitive fusion</td>
<td>Per target</td>
<td>NR</td>
<td>47</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Kuru et al [25]</td>
<td>1: Artemis</td>
<td>First software registration</td>
<td>Per target</td>
<td>53</td>
<td>32</td>
<td>9.8</td>
<td>13.2</td>
<td>7.6</td>
<td>0</td>
</tr>
</tbody>
</table>

NR = not reported.

In this study, the targeted approach, including visual and magnetic resonance-transrectal ultrasound fusion biopsy, is given as a whole; therefore, it was not possible to distinguish the outcomes of each targeted approach.

However, two studies were scored as potentially biased. One had a small sample size ($n = 13$) and significant patient heterogeneity [27]; the other had significant limitations, such as a retrospective design and patient heterogeneity [32].

In four studies, the index test domain demonstrated a high risk of bias and applicability concern. In two studies, this was due to an unclear interpretation of sampling results that was not clarified despite contacting the study authors [27,32]. In another study, the index test was correctly described and conducted, but the histologic outcomes of the image fusion targeted biopsies were partially pooled with the standard biopsy outcomes. Consequently, the comparative interpretation of results was not possible [26]. Finally, another study pooled together the results of both MRI-TRUS fusion biopsies and visual registration biopsy. Visual registration involves the biopsy operator making a judgement about where to deploy the needle by looking at the mpMRI on a separate screen [24].

None of the studies used an adequate reference test with 13 using the current standard of TRUS biopsy. One study used a more accurate test, namely transperineal biopsy using a limited sampling template protocol (not every 5 mm as often done for template mapping biopsies) [40], but there were concerns about incorporation bias. Systematic transperineal template cores were not taken from areas that were previously sampled using MRI-TRUS image fusion targeted biopsies [26]. As a result, the detection rate of clinically significant disease was given as a whole and thus comparison between the two could not be made.

In terms of flow and timing, one study was at high risk of bias because there was no adequate description and not all patients underwent the reference test [32]; two other papers did not describe flow and timing [20,27].

#### 3.2. Description of devices

In the studies included in this review, we identified eight image fusion platforms currently being used clinically to perform MRI-TRUS targeted biopsies (Supplementary Table 3). MRI-TRUS image fusion is most simply described as a way to align a preprocedure MRI to an intraprocedure US to accurately direct the biopsy needle to a US region of interest defined by mpMRI. Multiple devices available to do this can broadly be classified by the (1) fusion process as either rigid or nonrigid; (2) the mode of acquisition of the prostate US, either automatic or manual; and (3) the biopsy approach, either transrectal or transperineal.

Rigid fusion uses a direct overlay of the mpMRI onto the US images at the time of biopsy (Fig. 4). It does not take into account changes in shape and position of the prostate at the time of biopsy or try to compensate for this. Nonrigid fusion aims to compensate for prostate deformation at the time of biopsy. This occurs due to different positions in which the mpMRI and US are performed (supine vs left lateral or lithotomy), and the presence and movement of an endorectal US probe that causes changes in prostate shape.

The nonrigid fusion process can be achieved with elastic registration of the prostate US and MRI surfaces [41] or by
Fig. 2 – Independent risk of bias assessment per study using the Quality Assessment for Diagnostic Studies-2 tool. One paper reported two studies with identical design but using a different magnetic resonance-transrectal ultrasound fusion platform [20].

Fig. 3 – Summary of risk of bias assessment of all papers included using the Quality Assessment of Diagnostic Accuracy Studies-2 tool.
statistical motion modelling of how the prostate should deform using physical constraints [42].

3.3. Study designs

One paper reported two studies [20]; therefore 15 studies were included overall. Fourteen used a paired cohort design (Table 1). A total of 2293 men were included with a sample size ranging from 13 to 582. Three studies were conducted in biopsy-naïve men, three in men with a previous negative TRUS biopsy, eight studies reported on a mixed cohort who were either biopsy naive or had undergone a previous prostate biopsy, and one also included men with radio-recurrent disease (Table 1).

All mpMRI scans were performed on either a 1.5- or 3-T scanner. As a minimum, all men had T2-weighted scans and DWI. Fourteen of the 15 studies also performed DCE, three studies used MR spectroscopy (in addition to T2 weighting, diffusion, and dynamic contrast), and six studies required the addition of an endorectal coil. The mpMRIs were reported on a scale of suspicion, but the increments used differed between studies. For instance, the Likert scale is very similar to breast mammography reporting in which 1 confers the lowest radiologic suspicion for cancer and 5 means the radiologist is very confident that a lesion is cancer. Others used a 3- or 4-point scale, and a number reported mpMRI using the Prostate Imaging-Reporting and Data System that has a 5-point scale for each type of MRI sequence and then uses the aggregate total as the headline score (Table 1).

Fig. 4 – Description of the process of magnetic resonance (MR) to ultrasound image fusion. (a) The T2-weighted anatomic sequence is contoured (in red) to visualise the prostate. (b) Intraoperatively, the margins of the prostate are visualised by the surgeon and need to be contoured (in white) when using nonrigid registration. (c) Finally, in rigid fusion systems the surgeon overlaps manually the contouring of the prostate on the MR over the contour of the prostate on the ultrasound. In nonrigid fusion systems, the software tries to compensate for the difference in prostate contouring that is clearly seen as the area not aligned in this image.

The median detection of clinically significant disease was 23.6% (range: 4.8–52%) for standard biopsy and 33.3% (range: 13.2–50%) for MRI-TRUS image fusion targeted biopsy (Table 2). Across all studies in which both rates were reported, the use of MRI-TRUS fusion allowed the detection of greater numbers of clinically significant cancers compared with standard biopsy. The absolute difference in detection rate between the two approaches was a median of 6.8% (range: 0.9–41.4%) and always in favour of the approach based on MR-TRUS software.

Substantial discrepancy was found in the definition of clinically significant disease. Only one study did not report the criteria for defining this outcome, and there was no clarification provided by the authors when approached [23]. In all the remaining studies, the presence of Gleason pattern 4 was considered clinically significant disease. In eight studies, maximum cancer core length was also considered, although the threshold above which clinically significant disease was defined ranged from 3 mm to 10 mm.

3.5. Secondary outcomes

3.5.1. Detection of any cancer

The median detection rate of any cancer was 43.4% (range: 14.3–59%) and 50.5% (range: 23.7–82.1%) in the standard biopsy strategy versus MRI-TRUS image fusion biopsy, respectively. The absolute difference in overall detection of PCa between the two approaches was a median +6.9% in favour of the MRI-TRUS image fusion targeted biopsy approach (range: −8.8% to +53.2%). In four studies, standard biopsies detected more clinically insignificant disease than the software-based approach.

3.5.2. Efficiency

In all series, an image fusion approach was more efficient in detecting clinically significant disease. The median number of cores needed to detect one man with clinically significant cancer was 37.1 (interquartile range [IQR]: 32.6–82.8; range: 23.2–252) and 9.2 (IQR: 4.6–24.8; range: 4–37.7) for standard and MRI-TRUS image fusion targeted biopsy, respectively. The median difference in number of cores required across the series was 32.1 cores (IQR: 28.3–57; range: 21.4–84.8) in favour of the targeted approach. In other words, to detect the same number of clinically significant cancers with standard biopsy, approximately
four times the number of cores would be needed compared with an image fusion targeted approach.

3.5.3. Utility
Utility in our study was defined as the number of clinically significant cancers detected by one sampling strategy but missed by the other strategy. Considering absolute differences, MRI-TRUS image fusion biopsies detected a median of 9.1% additional clinically significant cancers (range: 5–16.2%) that were missed by standard biopsy alone. In contrast, standard biopsies detected a median of 2.1% (range: 0–12.4%) additional clinically significant cancers that were missed by MRI-TRUS fusion biopsies. However, if the study using transperineal mapping biopsies is removed so the standard biopsy is only a TRUS biopsy approach, the range stood at 0–7%.

3.5.4. Adverse events
Only two studies reported adverse events, but in none was a standard classification system used. One study stated that no adverse events occurred [25]. In another, four serious adverse events occurred, although it was not possible to distinguish which of the two approaches led to the adverse event because standard and targeted biopsies were performed in the same session under general anaesthesia [26].

3.6. Comparison of different magnetic resonance imaging-transrectal ultrasound image fusion targeted approaches
Two studies evaluated the outcomes of MRI-TRUS image fusion biopsies versus visual registration targeted biopsy (Table 3). One study did not report sufficient information to determine the primary outcome measure and a number of the secondary outcome measures. In the only outcome reported, namely detection of any cancer, MRI-TRUS image fusion biopsies had a higher rate (53% vs 47%; no \( p \) value given) [24]. The other study evaluated the two targeting approaches in 125 men presenting with a total of 172 targets. In a per target analysis, MRI-TRUS image software biopsies detected more clinically significant cancers (20.3% vs 15.1%; \( p = 0.05 \)) and more cancer overall (32% vs 26.7%; \( p = 0.14 \)). It also had better efficiency compared with visual registration requiring 9.8 rather than 13.2 cores to detect one man with clinically significant cancer [25]. There was no additional utility in using visual registration targeting, whereas the image fusion approach detected 7.6% additional clinically significant cancers that would have been missed by the visual registration approach. However, the study was underpowered to show the demonstrated absolute difference in detection rate of approximately 5% because it was powered a priori to demonstrate a 15% difference in detection rate.

3.7. Discussion
Our systematic review shows that MRI-TRUS image fusion targeted biopsies detect more clinically significant cancers using fewer cores compared with standard biopsy techniques. Most studies also showed a higher detection of clinically insignificant cancer, although four studies demonstrated a lower detection rate of clinically insignificant cancer by MRI-TRUS image fusion biopsies.

Our systematic review had some limitations. First, the definition of clinically significant disease varied between studies. This corresponds to the current uncertainty of identifying the determinants of cancer progression on biopsy. Although Gleason grade is an accepted prognostic indicator [43], it has been difficult to define the volume of cancer that is significant due to the current inaccuracies of the diagnostic pathway. Surrogate markers of cancer volume using cancer core length and number of positive biopsies are currently used [44]. The exact threshold for these is not clear. Historically, investigators have used lesion volume of 0.2 ml [45] or 0.5 ml [46] to define the threshold, although more recently, it was suggested that a pure Gleason 6 lesion could be 1.3 ml before defining it as clinically significant [47].

Second, it was not possible to determine the overall accuracy of a MRI-TRUS image fusion approach because an accurate reference test was not used. Therefore, although an image fusion approach for targeting seems to be superior to a standard TRUS biopsy, the residual diagnostic error of such an approach is not known. The UK Prostate Imaging Compared to Transperineal Ultrasound Guided Biopsy for Significant Prostate Cancer Risk Evaluation (PICTURE) study will evaluate this [48].

Third, most studies included a heterogeneous population. For instance, in men with previous negative findings or in men with low-risk disease on standard biopsy, applying another test may introduce selection bias. Although we agree this is possible, we would argue that given the consistency of the results across the studies, heterogeneity increases external validity and suggests wide application of MRI-TRUS image fusion in the spectrum of men looking for precise risk attribution. Particularly, the value of the fusion approach was consistently demonstrated in studies including only biopsy-naive men.

Fourth, the procedure of MR-TRUS fusion targeted biopsy is not standardised. There was significant variability across the studies in terms of MR characteristics and interpretation, threshold for biopsy, targeted biopsy conduct, and number of cores per target. Although this is to be expected during the development phase of a new interventional procedure, standardisation of the technique will allow better comparison between different software and against standard tests. Finally, only one study compared an image fusion approach to visual registration; further comparative research is needed in this area [48].

Despite these limitations, we believe our review has significant implications. First, when high-quality mpMRI is available, an image fusion approach might be offered in addition to standard sampling. This seems to allow better risk stratification with additional limited sampling and could therefore help improve decision making for men diagnosed with PCa. Second, despite the wide differences in software characteristics that indicate a possible difference in accuracy, this systematic review shows that the
outcomes are similar. Therefore the choice for one unit should be guided by factors such as cost, usability, interface, and flexibility of use. Third, this systematic review should guide future research. A more robust assessment of the diagnostic accuracy of image fusion targeted biopsy compared with an accurate reference test and at the same time with the current standard of practice, namely TRUS biopsy, is needed. This will require paired cohort studies or RCTs that have sufficiently large sample sizes with homogeneous populations to detect the differences we have reported here. These comparative studies need to be multicentre to ascertain validity and reproducibility. Equally, additional evidence is needed to compare visual registration with an image fusion because there are cost implications if new capital has to be purchased. Some of these trials are already under way, and the results will be available in the next few years [48].

The last aspect to consider is whether we are ready to change practice based on this review. There are a number of barriers to recommend immediate change. MRI-TRUS image fusion targeted biopsy needs to be considered a “complex procedure” [49], in which the final outcome depends on a chain of events from image acquisition, image interpretation by experts in the field, software accuracy, and finally biopsy operator ability, skill, and expertise. Each of these steps is important for the following one to be successful and therefore for the final outcome. This means that before recommending wide adoption, these elements need to be available on site, quality assured, and quality controlled so as not to affect the ultimate outcomes. MRI-TRUS image fusion might overcome the barriers inherent when using visual registration targeted biopsies because a high degree of expertise is required in specialist centres for the latter. With more than 1 million biopsies of the prostate performed every year in the United States and another estimated million in Europe, it will not be possible to centralise these biopsies in specialist centres. Therefore specific training in tertiary centres, courses at international and national meetings, and structured mentorship should be set up. These quality and training issues have been dealt with by programmes in breast cancer screening with mammography or breast MRI in high-risk groups [50].

Another significant issue are the cost and time implications. An image fusion approach combines the cost of the mpMRI plus the software/device. Although at face value, this seems more costly than a standard approach, the implications of potentially increased diagnostic accuracy and the downstream implications on follow-up and treatment should be taken into account. Early evidence shows that a targeted biopsy approach using MRI-TRUS fusion might be cost effective [51,52]. Time issues also need to be considered because MRI-TRUS fusion biopsies require additional time for contouring the lesions before the procedure and for image fusion and synchronisation. Although this added time might be problematic in terms of efficiencies in working, the added diagnostic value conferred by this targeted strategy might offset such additional resource implications.

4. Conclusions

Our systematic review shows that mpMRI-to-US image fusion targeted prostate biopsies detect more clinically significant cancers using fewer cores compared with standard biopsy techniques. Some studies confirmed a lower detection rate of clinically insignificant cancer using mpMRI to target the biopsies. Before this approach is incorporated into standard practice across all centres, consideration must be given to the need for the required expertise and skills and the current impediments to wider dissemination. If our findings were confirmed by large multicentre validating studies and also shown to be cost effective, MRI-TRUS image fusion targeted biopsies should be incorporated into the standard diagnostic pathway.

**Author contributions:** Massimo Valerio had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Valerio, Ahmed.

**Acquisition of data:** Valerio, Donaldson.

**Analysis and interpretation of data:** Valerio, Donaldson, Emberton, Ehdai, Hadaschik, Marks, Mozer, Rastinehad, Ahmed.

**Drafting of the manuscript:** Valerio, Donaldson, Ahmed.

**Critical revision of the manuscript for important intellectual content:** Emberton, Ehdai, Hadaschik, Marks, Mozer, Rastinehad, Ahmed.

**Statistical analysis:** Valerio, Donaldson, Ahmed.

**Obtaining funding:** None.

**Administrative, technical, or material support:** None.

**Supervision:** Emberton, Ehdai, Hadaschik, Marks, Mozer, Rastinehad, Ahmed.

**Other (specify):** None.

**Financial disclosures:** Massimo Valerio certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Massimo Valerio has received funding for conference attendance from Geoscan Medical. Boris A. Hadaschik has received research funding from Medcom and UroMed and is a paid consultant to Dendreon, Janssen, and Teva. He has also received funding for conference attendance from Bayer and Astellas. Pierre Mozer has been involved in the patent of the Urostation (Koelis) system. Mark Emberton and Hashim U. Ahmed receive funding from USHIFU, GSK, AngioDynamics, and Advanced Medical Diagnostics for clinical trials. Mark Emberton is a paid consultant to AngioDynamics, Steba Biotech, and SonaCare Medical (previously called USHIFU). Both have previously received consultancy payments from Oncura/GE Healthcare and Steba Biotech. None of these sources had any input regarding this paper. The SICPA Foundation supports the ongoing fellowship and PhD programme of Valerio Massimo. Ian Donaldson receives funding from the Wellcome Trust and Department of Health. Behfar Hadaschik receives funding from the German Research Foundation, German Cancer Aid, and the European Foundation for Urology. Mark Emberton and Hashim U. Ahmed would like to acknowledge funding from the Medical Research Council (UK), the Pelican Cancer Foundation charity, Prostate Cancer UK, St. Peters Trust charity, Prostate Cancer Research Centre, the Wellcome Trust, National Institute of Health Research-Health Technology Assessment programme, and the US National Institute of Health-National Cancer Institute. Mark Emberton receives funding in part from the UK National Institute of Health Research UCLH/UCL Comprehensive Biomedical Research Centre.
Leonard S. Marks was supported in his work at UCLA by Award Number R01CA158627 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

**Funding/Support and role of the sponsor:** None.

**Appendix A. Supplementary data**

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2014.10.026.

**References**


Epstein JL, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. JAMA 1994;271:368–74.


