Prostate-specific antigen vs. magnetic resonance imaging parameters for assessing oncological outcomes after high intensity–focused ultrasound focal therapy for localized prostate cancer

Louise Dickinson, M.B.B.S., Ph.D.\(^a,b,1,*\), Hashim U. Ahmed, Ph.D., F.R.C.S. (Urol)\(^a,b\), Richard G. Hindley, F.R.C.S. (Urol)\(^c\), Neil McCartan, M.Sc.\(^a\), Alex Freeman, F.R.C. Path\(^d\), Clare Allen, F.R.C.R.\(^e\), Mark Emberton, M.D., F.R.C.S. (Urol)\(^a,b\), Alex P. Kirkham, F.R.C.R.\(^e\)

\(^a\) Division of Surgery and Interventional Sciences, University College London, London, UK
\(^b\) Department of Urology, University College London Hospitals NHS Foundation Trust, London, UK
\(^c\) Department of Urology, Basingstoke Hospital, Hampshire Hospitals NHS Foundation Trust, Hampshire, UK
\(^d\) Department of Histopathology, University College London Hospitals NHS Foundation Trust, London, UK
\(^e\) Department of Radiology, University College London Hospitals NHS Foundation Trust, London, UK

Received 16 May 2016; received in revised form 11 July 2016; accepted 21 July 2016

Abstract

Introduction: Focal therapy for localized prostate cancer has the potential for oncological control without the side effects of radical therapies. However, there is currently no validated method for monitoring treatment success. We assessed the diagnostic performance of prostate-specific antigen (PSA) parameters and MRI compared to histological outcomes following focal therapy.

Patients and methods: Patients from 3 Ethics Review Board approved prospective studies of focal high intensity–focused ultrasound (HIFU) (Sonablate 500) for localized prostate cancer (T1c–T3a, Gleason grade \(\leq 4 + 3\), and PSA \(\leq 20\)). Post-HIFU PSA nadir, 6-month PSA, PSA density, and early ( \(< 3\) wk) and late (6 mo) MRI (T2-weighted, dynamic contrast-enhanced \(\pm\) diffusion-weighted) was assessed for predictive accuracy of cancer on postoperative biopsy, using receiver operating characteristic (ROC) analysis and sensitivity, specificity, and positive and negative predictive estimates. ROC areas for MRI and PSA were compared. Calculations for statistical significance (\(P \leq 0.05\)) were obtained in a subset of patients comparing area under ROC for 6-month MRI and PSA criteria, across 4 different histological definitions of disease significance.

Results: Of 118 men, 111 underwent at least 1 postoperative biopsy (median 6 cores), with an overall positive biopsy rate of 37% (41/118), over a mean follow-up period of 716 days post-HIFU. Areas under ROC for early and late MRI were (depending on definition of significant disease) 0.65 to 0.76 and 0.77 to 0.85, respectively, with sensitivity, specificity, and negative predictive values of 68% to 91%, 52% to 55%, and 85% to 98% (early MRI), and 63% to 80%, 67% to 73%, and 86% to 97% (late MRI). The area under the ROC curve was statistically significantly higher for late MRI than 6 months and nadir PSA for residual disease > 3 mm or any Gleason 4 tumor.

Conclusions: Early and late MRI performed better than PSA measurements in the detection of residual tumor after focal therapy.

Keywords: Focal therapy; Follow-up; MRI; PSA; HIFU

1. Introduction

Focal therapy is an emerging treatment for localized prostate cancer, using a range of technologies [1–6]. Early reports of high intensity–focused ultrasound (HIFU) and cryotherapy have demonstrated low rates of genitourinary side effects, in conjunction with acceptable short-term cancer control [7]. However, one of the principal concerns in focal therapy is the lack of a validated, noninvasive test for monitoring oncological outcome [8]. Although prostate-specific antigen (PSA) parameters are an established means of monitoring biochemical outcomes after...
Biopsy is invasive, with an associated risk of significant morbidity [9], and we know that systematic biopsies for small foci of tumor are subject to considerable undersampling and undergrading compared to targeted samples [10]. Conversely, imaging has the capability of visually monitoring the whole prostate (treated and untreated areas) and of providing information on changes in characteristics that might signal residual disease, progression of untreated secondary lesions, or development of de novo lesions. Magnetic resonance imaging (MRI) has also shown good performance for the detection of residual tumor after whole prostate HIFU ablation, with sensitivities for the detection of residual disease of 73% to 87% [11].

We aimed to assess the diagnostic performance of PSA parameters and multiparametric MRI (mpMRI) for the detection of residual disease found at biopsy after focal HIFU, using 4 histological threshold definitions of significant disease.

2. Patients and methods

2.1. Patients

A total of 3 early development studies evaluating focal HIFU treatment (Sonablate 500) for localized prostate cancer have now been completed at our institutions. Research committee approval and written, individual patient consent was obtained for each study. In the first study \( n = 20 \) [1], treatment was delivered as a hemiablation to unilateral disease; in the second study \( n = 42 \) [2], treatment involved more targeted “focal” ablation of cancer lesion(s); and in the third \( n = 56 \), “index lesion” ablation was performed to the dominant clinically significant lesion(s) only [3]. The HIFU treatments occurred between 2006 and February 2011, and all men entered a prospective HIFU registry on completion of the study period for collection of ongoing oncological (including PSA, mpMRI, and biopsy) and quality-of-life data. All of the 118 patients in these trials were included.

Short-term (12 mo) adverse event and genitourinary functional, biochemical, and histological outcomes have been published on the results on all 3 studies [1–3]. The purpose of this current report was to assess the accuracy of MRI and PSA outcomes against histopathology on all participants within the 3 studies who underwent at least 1 post-HIFU biopsy. Additionally, this analysis includes registry data extending beyond the 12-month outcomes previously published [1–3].

2.2. Magnetic resonance imaging

All pre-HIFU and post-HIFU MRI were performed using either a 1.5 or 3-T MRI scanner and a pelvic phased array receiver, pelvic coil, full protocol of T1- and T2-weighted turbo-spin echo images and dynamic postgadolinium volume acquisition. Diffusion-weighted images were also obtained in all patients pre-HIFU and in 15 patients post-HIFU (6-mo scan) (Appendix Table A).

The early MRI was performed between 48 hours and 3 weeks post-HIFU, as a preliminary assessment of treatment quality, with coverage of the tumor scored between 1 and 3 (1 indicating the lowest suspicion of residual tumor). If the margin of the nonenhancing prostate was close to the treated tumor, we scored 2, and if there was enhancement in the tumor location, we scored 3.

The late MRI was performed at approximately 6 months, pre-follow-up biopsy. This was the earliest stage that evidence of any residual tumor was expected to become apparent. Before this time, resolving necrotic tissue has been found to mask residual cancer [12]. Early and late MRI were reported and graded prospectively by one of a group of specialist uroradiologists with at least 3 years of prostate MRI experience. A 5-point Likert-type scale for likelihood of residual tumor was used, with a score of 3 or higher considered positive (as per standard departmental practice, before the introduction of the PI-RADS scoring system). In a few cases where no score had been assigned prospectively, the images were reviewed and graded retrospectively by a single experienced uroradiologist (AK), without knowledge of the posttreatment histological outcome (Fig. 1).

2.3. Prostate-specific antigen

Serum PSA levels were obtained pre-HIFU, and at 6 weeks, 3, 6, 9, and 12 months, and approximately 6-monthly thereafter. PSA nadir was calculated as the lowest PSA achieved at any point following focal HIFU. Postoperative PSA density was calculated as 6-month PSA level divided by the volume of residual tissue, as assessed by the 6-month post-HIFU MRI.

2.4. Biopsies

A limited number of representative biopsies were taken at 6 months from the treated area of the prostate, at an approximate density of 1 per milliliter of residual tissue as assessed on ultrasound. If the late MRI indicated any areas of suspicion, these were also targeted at biopsy (cognitive targeting). This protocol was standardised across all of the focal therapy studies included in this analysis. Additionally, “for-cause” biopsies were performed in some men with ongoing suspicion of residual tumor in later follow-up (either a suspicious MRI or continuing rise in PSA). Only biopsies in the area of the treatment were included in our analysis.
2.5. Disease stratification

In order to differentiate between different burdens of residual tumor, and to assess thresholds for detection by imaging and biochemical parameters, we divided post-HIFU tumor into the following categories:

1. Any cancer.
2. Gleason pattern 4 or 5, or >3 mm maximum cancer core length, or both (UCL definition 2) [13].
3. A >3 mm maximum cancer core length (any disease grade).
4. Gleason pattern 4 or 5 (any cancer core length).

2.6. Statistical analysis

PSA nadir, 6-month postoperative PSA, postoperative PSA density, and early and late MRI outcomes were assessed for predictive accuracy of histological residual tumor on postoperative biopsy using area under the receiver operating characteristic (AUROC) analysis. All figures given are for fitted curves produced with the online ROC calculator ROCfit, which generates the maximum likelihood fit of a binormal model [14]. We performed the following 2 analyses: firstly using only the 6-month biopsy data, and secondly including the result of any positive biopsy. Areas under the ROC curve for MRI and the PSA criteria were compared using the method described by Hanley and McNeil [15] (Fig. 2).

Additionally, estimates of sensitivity, specificity, and positive predictive values (PPV) and negative PV (NPV) were calculated after dichotomising the MRI results so that equivocal scans were considered positive (2 or 3 considered positive on the early scan, and 3–5 considered positive at 6 mo)

3. Results

Of 118 men with localized prostate cancer (T1c–T3a, Gleason grade ≤4 + 3, and PSA ≤ 20) who underwent focal ablation within any of the 3 prospective 12-month

![MRI images](image-url)

Fig. 1. MRI images before (A and B), early post (C), and 6 months after focal HIFU to a Gleason 3 + 4 right anterior lesion. The lesion is shown on T2 (A) and dynamically enhanced (B) images (arrow in each) before treatment. In early post-HIFU (C), the treatment margin is close to the location of the tumor (arrow), and the score was equivocal for residual tumor (score 2). Late post-HIFU (D), a rim of enhancement at the site of the tumor was scored as positive on MRI (score 5), and the biopsy was positive for all definitions of significance.
studies, 111 received at least 1 postoperative biopsy (at a median of 190 days post-HIFU, with a median of 6 cores). Of all, 16 men received a second “for-cause” biopsy (a mean 716 days post-HIFU), due to suspected residual or recurrent tumor (usually a suspicious MRI or a rise in PSA). Baseline characteristics are detailed in Appendix Table B.

Of the 111 men, 109 with a 6-month biopsy also had a 6-month MRI and full PSA data. This group was used to analyze both the performance of 6-month MRI and PSA measurements. The performance of early MRI was assessed using a smaller group of 105 men who underwent a scan at 0 month and biopsy at 6 months.

3.1. Biopsy outcomes

Overall, 28 men (25%) had residual disease on first biopsy, of whom 12/28 (43%) fulfilled our criteria 2, 3, or 4 for clinically significant disease. After for-cause biopsy, a total of 41 men (37%) had received a positive biopsy posttreatment, of whom 21/41 (51%) fulfilled criteria 2, 3, or 4 for clinical disease significance. The median cancer core length of residual disease on any biopsy was 2 mm (range: 1–10 mm).

3.2. PSA parameters

The median pre-HIFU PSA was 7 ng/ml, and median posttreatment nadir, 6-month, and PSA density values were 1.6, 2.0, and 4.5 ng/ml, respectively. A 6-month and nadir PSA levels demonstrated similar AUROC accuracies for residual disease detection against initial biopsy (0.63–0.71).

For both groups, additional information on subsequent biopsy resulted in globally reduced accuracy rates. The results for PSA density were similar, but with a higher accuracy (0.78) for the detection of Gleason 4 or 5 disease (Tables 1 and 3).

The P values for the comparison of the PSA parameters AUC with that for 6-month MRI (Table 4) are given in Appendix Table C.

3.3. Magnetic resonance imaging

3.3.1. Early MRI

A total of 105 men received an early MRI (at a median of 15 days after biopsy, range: 7–29) of which 54 (51%) were considered equivocal or positive. Moderate accuracy rates were achieved for predicting the presence of residual cancer, with AUROC ranging between 0.65 and 0.76. NPV for residual tumor was high (85% for any tumor, but 98% for the 3 definitions of significant disease), but PPVs were low across all disease categories (range: 9.4%–32.1%).

When early MRI outputs were correlated with the outcomes of any positive biopsy, including repeat biopsy in those 16 men receiving one, sensitivity and NPVs decreased marginally for the detection of clinically significant tumor. However, specificity rates remained similar, and PPVs improved.

3.3.2. Late MRI (6-mo follow-up)

Of 109 men, 38 (35%) were considered equivocal or positive on MRI at a median of 6 months (range: 154–280 d). Against initial biopsy, late MRI demonstrated greater accuracy for predicting residual tumor than early MRI, with AUROC ranging between 0.77 and 0.85 (highest for disease >3 mm). As with the early MRI, the late MRI demonstrated high NPVs for ruling-out the presence of clinically significant tumor (86%–97%) but low PPVs (14%–44%). Specificity rates were similar across all disease categories.

Again, the additional information from subsequent positive biopsy contributed only marginally, with the greatest change observed in the PPVs of late MRI for ruling-in residual cancer.

A summary of the postoperative PSA and MRI outcomes is detailed in Appendix Table D.
3.4. Statistical comparisons

To avoid effects owing to large numbers of comparisons, we restricted the statistical tests to comparing 6-month MRI with PSA criteria in a single, defined group of 109 patients. In all 24 comparisons (P values given in Table 2), the AUC was higher for MRI than for PSA criteria, and in 6 cases the difference was statistically significant.

4. Discussion

In this short-term study, MR imaging provided a more accurate means of detecting residual cancer than PSA biochemical parameters, against a biopsy histopathology reference standard, with high AUROC of up to 0.85 by late MRI in the detection of different criteria for clinically significant tumor.

Early and late MRI achieved very high NPVs, at 98% and 97%, respectively, for ruling-out the presence of residual clinically significant tumor on posttreatment biopsy. NPV in this context must be interpreted with caution, as the number of positive biopsies was low for some of the definitions of significance. The low PPVs achieved are likely to reflect the chosen cutoff of 3 on the Likert scale as “positive,” and threshold effects due to small tumors being classified as “negative” for some of the definitions of significance.

Our results suggest that mpMRI is a useful test for assessing focal HIFU treatment, and in particular, for confirming the absence of residual cancer in the short term. Although the NPVs of the early and later MRIs are similar, the considerably higher specificity and PPV of the later (6 mo) scan means that it is potentially more useful.

The clinical implications extend beyond HIFU ablation and are potentially applicable to any focal therapy. Firstly, those assessed as at high risk for undertreatment could be identified early on MRI, and closely monitored, or offered further treatment at the first appropriate opportunity. Secondly, MRI may enable a reduction in the number of cores performed during follow-up biopsies by targeting positive areas, or a reduction in the need for posttreatment biopsy if negative, although this latter application remains speculative.

By analyzing accuracy rates against several definitions of residual tumor burden, we were able to show that the highest accuracy levels were achieved when identifying lesions more than 3 mm. The term “clinically significant” has recently been adopted within research papers in order to

| Table 2 | ROC characteristics of PSA parameters against initial and subsequent biopsy |
|-----------------|-----------------|-----------------|-----------------|
|                | Any disease     | > 3 mm CCL or   | > 3 mm CCL      | Any Gleason      |
|                |                 | any Gleason 4   |                 | pattern 4       |
| **Initial biopsy (standard errors in brackets)** |                  |                 |                 |                 |
| PSA nadir      | 0.63 (0.055)    | 0.64 (0.072)    | 0.71 (0.066)    | 0.66 (0.096)    |
| 6-month PSA    | 0.64 (0.052)    | 0.65 (0.068)    | 0.71 (0.058)    | 0.67 (0.092)    |
| PSA density    | 0.64 (0.070)    | 0.66 (0.070)    | 0.67 (0.077)    | 0.78 (0.064)    |
| **Any biopsy (standard errors in brackets)**  |                  |                 |                 |                 |
| PSA Nadir      | 0.58 (0.053)    | 0.56 (0.064)    | 0.67 (0.070)    | 0.54 (0.077)    |
| 6-month PSA    | 0.58 (0.053)    | 0.56 (0.065)    | 0.67 (0.066)    | 0.50 (0.081)    |
| PSA density    | 0.62 (0.054)    | 0.59 (0.073)    | 0.67 (0.082)    | 0.59 (0.093)    |
| CCL = cancer core length. |

| Table 3 | Accuracy rates of early MRI against initial and subsequent biopsy |
|-----------------|-----------------|-----------------|-----------------|
|                | Any disease     | > 3 mm CCL or   | > 3 mm CCL      | Any Gleason 4   |
|                |                 | any Gleason 4   |                 | pattern 4       |
| **Initial biopsy** |                  |                 |                 |                 |
| Sensitivity (%) | 17/25 (68%)     | 10/11 (91%)     | 9/10 (90%)      | 5/6 (83%)       |
| Specificity (%) | 44/80 (55%)     | 51/94 (54%)     | 51/95 (54%)     | 51/99 (52%)     |
| PPV (%)         | 17/53 (32%)     | 10/53 (19%)     | 9/53 (17%)      | 5/53 (10%)      |
| NPV (%)         | 44/52 (85%)     | 51/52 (98%)     | 51/52 (98%)     | 51/52 (98%)     |
| ROC (standard error) | 0.65 (0.070) | 0.76 (0.063)    | 0.76 (0.067)    | 0.75 (0.11)     |
| **Any biopsy**  |                  |                 |                 |                 |
| Sensitivity (%) | 25/37 (68%)     | 15/19 (79%)     | 12/14 (86%)     | 8/12 (67%)      |
| Specificity (%) | 40/68 (59%)     | 48/86 (56%)     | 50/91 (55%)     | 48/93 (52%)     |
| PPV (%)         | 25/53 (47%)     | 15/53 (28%)     | 12/53 (23%)     | 8/53 (15%)      |
| NPV (%)         | 40/52 (77%)     | 48/52 (92%)     | 50/52 (96%)     | 48/52 (92%)     |
| ROC (standard error) | 0.69 (0.064) | 0.72 (0.063)    | 0.75 (0.064)    | 0.64 (0.10)     |
| CCL = cancer core length. |
Table 4
Accuracy rates of late MRI against initial and subsequent biopsy

<table>
<thead>
<tr>
<th></th>
<th>Any disease</th>
<th>&gt;3 mm CCL or any Gleason 4</th>
<th>&gt;3 mm CCL</th>
<th>Any Gleason pattern 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial biopsy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>17/27 (63%)</td>
<td>9/12 (75%)</td>
<td>8/10 (80%)</td>
<td>5/7 (71%)</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>60/82 (73%)</td>
<td>67/97 (69%)</td>
<td>68/99 (69%)</td>
<td>68/102 (67%)</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>17/39 (44%)</td>
<td>9/39 (23%)</td>
<td>8/39 (21%)</td>
<td>5/39 (13%)</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>60/70 (86%)</td>
<td>67/70 (96%)</td>
<td>68/70 (97%)</td>
<td>68/70 (97%)</td>
</tr>
<tr>
<td>ROC (standard error)</td>
<td>0.77 (0.054)</td>
<td>0.83 (0.056)</td>
<td>0.85 (0.053)</td>
<td>0.79 (0.081)</td>
</tr>
<tr>
<td><strong>Any biopsy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>24/40 (60%)</td>
<td>14/21 (67%)</td>
<td>11/14 (79%)</td>
<td>9/14 (64%)</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>54/69 (78%)</td>
<td>63/88 (72%)</td>
<td>67/95 (71%)</td>
<td>65/95 (68%)</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>24/39 (62%)</td>
<td>14/39 (36%)</td>
<td>11/39 (28%)</td>
<td>9/39 (23%)</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>54/70 (77%)</td>
<td>63/70 (90%)</td>
<td>67/70 (96%)</td>
<td>65/70 (93%)</td>
</tr>
<tr>
<td>ROC (standard error)</td>
<td>0.76 (0.068)</td>
<td>0.78 (0.067)</td>
<td>0.85 (0.056)</td>
<td>0.72 (0.087)</td>
</tr>
</tbody>
</table>

CCL = cancer core length.

differentiate between potentially indolent tumor compared to higher-risk disease that is likely to warrant treatment [16–19], although none have yet been validated. For this study, we adopted several criteria of disease burden, using 3 mm and Gleason 4 as thresholds of clinical significance according to UCL criteria 2 [13], not for the purpose of establishing a definition for clinical significance, but rather to assess different thresholds of volume and grade that might be predicted by either biochemical or imaging tools.

We should emphasize that the diagnostic performance of MRI in the postfocal HIFU context appears potentially better than in the more common setting of a triage test to investigate men with a raised PSA—only 10 out of our 28 positive patients had a cancer core length >3 mm, and our sensitivity of 80% for a biopsy with a cancer core length >3 mm compares well with the finding of Villers et al. of a sensitivity of 77% for lesions >0.2 cc in the prebiopsy setting [20] (a 0.2 cc lesion is equivalent to a 7 mm sphere, and usually returns a maximum cancer core length of >3 mm [13]).

There are many limitations to this study. Although the cohort had all received focal therapy within three 12-month prospective studies involving protocol-mandated serum PSA, MRI, and biopsies, and subsequently prospective registry data collection, our analysis included some MRI data collated in a retrospective manner, albeit blinded to the postoperative PSA and histology results.

The analysis was also performed on a heterogeneous cohort of men, with varied baseline disease burden, undergoing different focal therapy protocols. Of particular relevance, the postoperative PSA response in those men treated with ablation to the index lesion and with known untreated clinically insignificant disease may have been significantly different to those without untreated lesions. However, it could be argued that all the treatment protocols adopted an “index lesion” approach, as the complete absence of low volume, and insignificant disease within the untreated prostate cannot be fully excluded on preoperative mpMRI and transperineal template mapping biopsies.

All of the focal therapy studies included were initiated before contemporary consensus statements of patient eligibility and follow-up criteria [21]. As men received the first postoperative biopsy within a relatively short follow-up period (averaging just over 7 mo from treatment), biopsy may not have accurately targeted and sampled any small residual lesions, potentially introducing sampling bias. It was in order to capture such cases that we included subsequent biopsy results, although with no significant effect identified on the PSA and MRI performance characteristics.

Finally, the MRI protocol changed at our institution over the study period of 6 years. Although most patients underwent 1.5 T scans involving just T2 and dynamically enhanced sequences, some had 3 T scans, and others also underwent diffusion-weighted scans. Both of these are likely to improve performance (with some data on diffusion existing in the posttreatment context [22]) so that we may be underestimating the potential performance of modern MRI.

5. Conclusions

Although clinical guidance now exists on the most accurate means of localizing and identifying disease to plan and conduct focal treatment, our understanding of how men receiving focal therapy are most appropriately followed up remains limited. This small, short-term study provides a first step to reducing that uncertainty. We have shown that MR imaging, in the form of early and later mpMRI, strongly predicts a negative biopsy after focal therapy for localized prostate cancer, while PSA parameters are less reliable.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.urolonc.2016.07.015.
References


