Abstract

Introduction: We aimed to report the clinically significant prostate cancer (PCa) detection rate in men undergoing magnetic resonance imaging-transrectal ultrasound (MRI-TRUS)-cognitive fusion (CF) targeted biopsies stratified by the Prostate Imaging and Data Reporting System (PI-RADS) version 2 (v2) scores.

Methods: With a quality assurance waiver from the IRB, we identified a cohort of men who underwent MRI-TRUS-CF and synchronous template biopsy from 2015–2017. MRI (PI-RADS v2 score, lesion size, lesion location [peripheral or transition zone (PZ/TZ)]), and CF-TRUS biopsy (operator experience, TRUS visibility, and number of biopsies) features were extracted. The primary outcome was diagnosis of clinically significant (Gleason score ≥3+4=7 or International Society of Urological Pathology (ISUP) grade group ≥2) PCa.

Results: During the study period, 131 men (with 142 PI-RADS v2 score ≥3 lesions) met inclusion criteria; 98 men had previously negative template biopsy and 33 were on active surveillance for previously detected low-grade PCa. In total, 41.9% (55/131) men had clinically significant PCa — 17.6% (23/131) detected on targeted biopsy only, 8.4% (11/131) on template biopsy only, and 16.0% (21/131) on both targeted and template biopsy. Clinically significant PCa detection stratified by PI-RADS v2 scores were: 11.1% (3/27) for score 3 (indeterminate), 42.9% (24/56) for score 4 (significant cancer likely), and 35.6% (21/59) for score 5 (significant cancer very likely). Clinically significant PCa detection rates in targeted biopsies were better among PZ (41.8% [33/79]) compared to TZ (23.8% [15/63]) lesions (p=0.025) in TRUS visible lesions (p=0.033) and in the most experienced radiologists (p=0.05), with no difference by lesion size or number of additional core biopsies performed (all p>0.05).
Conclusions: Cognitive fusion MRI-TRUS-guided targeted biopsy yielded substantially lower rates of clinically significant cancer in PI-RADS v2 score 4 and 5 lesions when compared to published results using in-bore MR-guided or automated MRI-TRUS fusion guidance systems. Cancer detection was worst for TZ lesions.

Introduction

Multi-parametric (mp) prostate MRI (the combination of anatomic T2-weighted imaging with functional imaging techniques [Diffusion weighted imaging and dynamic contrast enhanced MRI]) is the reference standard test for imaging the prostate 1. mp-MRI is accurate for detection of clinically significant prostate cancer (PCa), defined as Gleason score ≥ 3+4=7 tumor measuring ≥ 0.5 mL in size 2. mp-MRI is particularly valuable for detection of tumors located anterior to the urethra (in the anterior horns of the peripheral zone [PZ] and in the transition zone [TZ]); areas which are undersampled or not sampled at all during template transrectal ultrasound (TRUS) guided biopsy 3. mp-MRI is increasingly being used during active surveillance (AS) for men with low volume low grade (Gleason score 3+3=6) PCa with discordant clinical findings 4,5 and in men with previously negative template biopsies and high clinical suspicion of PCa 6. mp-MRI is also of value in biopsy naïve men, although the data in this patient population, particularly in Canada, remains immature 6.

While mp-MRI has been widely adopted by Urologists in Canada, is now a fundamental component of PCa management and improves diagnosis of cancers that may be occult on template biopsy; mp-MRI has also created new challenges. Lesions detected on mp-MRI often require targeted biopsy to be confirmed as histologically representing clinically significant PCa to better inform management decisions. Histological confirmation is helpful because false positive MRI interpretations can occur 7 and because not all lesions detected on mp-MRI have the same probability of representing significant PCa. In 2015, the Prostate Imaging and Data Reporting System (PI-RADS) version 2 (v2) document was released to standardize reporting of prostate mp-MRI and provide Urologists and other physicians managing PCa with a probability scale (from 1 to 5) of how likely a lesion detected on mp-MRI represents significant PCa 1,8. A PI-RADSv2 score of 4 or 5 typically warrants targeted biopsy due to “likely” and “very likely” probability scores of clinically significant PCa 1,8.

Targeted biopsy of mp-MRI detected lesions can be performed in-bore (i.e. within the MRI suite using MRI guidance), through advanced automated fusion of mp-MRI data onto real time 3-Dimensional (3D) TRUS images or through cognitive fusion of mp-MRI and 2D TRUS (where an operator mentally fuses MRI and TRUS images while performing biopsies) 9. A recent meta-analysis showed equivalent significant PCa detection rates comparing in-bore MR guided and automated MRI-TRUS fusion guided biopsies which both outperformed cognitive fusion; however, reporting of MRI among included studies pre-dated PI-RADSv2 10. In a previous Canadian study, Cool et al. demonstrated that automated MRI-TRUS fusion was superior to cognitive fusion for PCa detection even among experienced operators 11; however, similarly, this
study was published in the era before PI-RADSv2 which may have resulted in more false negative MRI interpretations and makes comparison to other studies reporting PCa detection with targeted biopsy difficult. The purpose of the present study is to report the rates of clinically significant PCa detection using mp-MRI-TRUS cognitive fusion biopsies in men with mp-MRI detected lesions stratified by PI-RADSv2 score from a single institution tertiary care referral center for prostate cancer and to evaluate factors which may influence significant PCa detection at cognitive fusion biopsy.

Methods

Patient selection and mp-MRI
With a quality assurance waiver from the IRB, we performed a search using our institutional Picture Archiving and Data Reporting System (PACS; Horizon Medical Imaging, McKesson corporation, San Francisco CA) to identify all patients who underwent cognitive fusion MRI-TRUS guided biopsy of the prostate between January 2015 and June 2017. We identified 236 patients and excluded 13 men in whom targeted biopsy was performed due to suspicion of locally recurrent tumor after radical prostatectomy or radiotherapy. Of the 223 remaining patients, all mp-MRI examinations were reviewed by an expert genitourinary (GU) radiologist (BLINDED) with 13 years of experience in prostate mp-MRI having interpreted over 500 prostate mp-MRI examinations using PI-RADSv2. Radiologist (BLINDED) also serves as the Director of Prostate Imaging at our institution (BLINDED). A review of each mp-MRI examination was conducted by (BLINDED) who was blinded to patient information including results from cognitive fusion biopsy and original MRI reports.

After dedicated review, 92 patients were further excluded because: mp-MRI was degraded by severe image artifact (N=4), examinations were re-interpreted as negative (PI-RADSv2 score 1 or 2, N=67), examinations were considered positive (PI-RADSv2 score ≥3) but in a discrepant location from the initial interpretation (with the new lesion identified after secondary review not having been sampled at time of biopsy, N=11) or examinations were performed at 1.5 Tesla due to a contraindication to imaging at 3 Tesla (N=10). From the 131 remaining patients, 142 lesions were identified with PI-RADSv2 assessment categories ≥3 that underwent targeted biopsy. PI-RADSv2 category, maximum size of lesion (measured on transverse axial T2W images) and location of lesions (peripheral zone [PZ; base, middle gland and apex] or transition zone [TZ]) were recorded. All mp-MRI examinations were performed on a clinical 3 Tesla MRI system (Discovery 750W, General Electric Medical, Milwaukee WI) using integrated body array coils (endorectal coil was not used) and with sequence parameters compliant with PI-RADS version 2.1,8 as described previously.12-17

mp-MRI TRUS-guided cognitive fusion-targeted biopsy and histopathology results
Targeted biopsies were performed using TRUS guidance with cognitive fusion of mp-MRI data onto real time 2-Dimensional TRUS images. All Ultrasound examinations were performed using
modern Ultrasound equipment (Aloka Prosound Alpha 10, Aloka Hitahi Medical or General Electric Logiq E9, General Electric Healthcare) using endoluminal 4-8 MHz end-fire probes. All biopsies were performed by fellowship trained abdominal Radiologists. During the time period of the study, 12 Radiologists performed at least 1 biopsy with a mean number of biopsies of 4 ± 3 (inter-quartile range 2 to 5). Our cognitive fusion biopsy program started in January of 2014; however, we included only patients who received biopsy in January 2015 and later to ensure there was an adequate learning period for Radiologists performing MRI-TRUS cognitive fusion biopsies and because standardized reporting of the cognitive fusion biopsy procedure (containing information used in this study) was instituted in 2015. Among the 12 Radiologists performing targeted biopsies, 5 Radiologists comprising the majority of the prostate biopsy service performed 91.5% (130/142) biopsies (range 16-37 biopsies per Radiologist) compared to the other 7 Radiologists who performed only 8.5% (12/142) biopsies (range 1-2 biopsies per Radiologist).

The TRUS-guided biopsy system used for all biopsies employed an 18-gauge side-cutting needle. All of our biopsy suites are equipped with monitors which enable display of mp-MRI which can be reviewed before or during the biopsy procedure. In 90.8% (119/131) men anesthesia was provided using 1% Lidocaine nerve block, whereas, in the remaining 9.2% (12/131) men anesthesia was provided using both a 1% Lidocaine nerve block in addition to conscious sedation. A fleet enema was prescribed prior to the procedure and antibiotic prophylaxis was prescribed to prevent infection as described previously 18. In 97.9% (128/131) men, at time of targeted biopsy a simultaneous extended sextant template biopsy of the PZ was also performed including 12 biopsies (two each from the bilateral base, middle and apical portions of the PZ) in accordance with provincial standards for PCa diagnosis. Core-needle biopsy specimens are submitted for laboratory processing and interpretation in separate pathology specimen containers according the site of sampling. Tissue from biopsy specimens are fixed overnight in 10% neutral buffered formalin. Three histological slides are prepared from each block, each with three serial sections cut at 3μm in thickness and stained with haematoxylin and eosin (H&E). Biopsy results are reported for each core specimen individually.

The cognitive fusion TRUS-guided biopsy reports used at our institution employ a mandatory standardized reporting template which specifies: the operator, the number of core biopsies performed per target and whether the mp-MRI detected target was visible on TRUS. These variables were recorded by an abdominal radiology fellow (BLINDED). The radiology fellow also retrieved the biopsy results from the patient electronic medical record. The presence of cancer at a biopsy site (targeted or template) was recorded by the radiology fellow (BLINDED) who also recorded the individual Gleason score from core-needle biopsy specimens at each biopsy site. A Gleason score of ≥ 3+4=7 was considered clinically significant in this study. In this way, a biopsy result for each targeted lesion, as well as, the remainder of the PZ for men who underwent template biopsy at time of fusion biopsy was recorded.
Statistical analysis
The proportion of detection of all cancers on a per patient and per lesion basis was tabulated; however, for all further comparisons only clinically significant cancers were considered. Clinically significant cancer detection rates were compared by lesion PI-RADSv2 score, size on MRI and location on MRI, as well as, operator experience, TRUS visibility and number of core biopsies performed per lesion. Comparisons were performed using independent t-tests, Chi-square and logistic regression for multi-variable analyses. Statistical analyses were performed using STATA version 13.0 (Statacorp, College Station TX) and p-values <0.05 were considered statistically significant outcomes.

Results
During the study period, 131 patients having 142 lesions with PI-RADSv2 assessment categories ≥3 who met the inclusion criteria underwent targeted biopsy. Mean patient age was 66 ± 8.2 (Range 48-86) years and mean prostate serum antigen (PSA) was 12.8 ± 11.3 (Range 2.1-64) ng/mL. In total, a diagnosis of any PCa, including Gleason score 3+3=6 tumors, was established in 74.8% (98/131) men. When excluding Gleason score 3+3=6 PCa, clinically significant cancers were diagnosed in 41.9% (55/131) of men; with 17.6% (23/131) diagnosed on targeted biopsy only, 8.4% (11/131) diagnosed on simultaneous template biopsies only and 16.0% (21/131) diagnosed on both targeted and template biopsies in the same patient.

Cancer detection rates on targeted biopsies stratified by PI-RADSv2 assessment category are provided in Table 1. From 142 lesions, PI-RADSv2 assessment categories were: score 3 in 19.0% (27/142), score 4 in 39.4% (56/142) and score 5 in 41.5% (59/142). The detection rates of any cancer, including Gleason score 3+3=6 cancers, by PI-RADSv2 scores were: 40.7% (11/27) score 3, 67.9% (38/56) score 4 and 69.5% (41/59) score 5. The detection rates of clinically significant cancer by PI-RADSv2 scores were: 11.1% (3/27) score 3, 42.9% (24/56) score 4 and 35.6% (21/59) score 5. There were significantly higher clinically significant cancers detected in score 4 and 5 lesions compared to score 3 lesions (p=0.015).

Lesions in the PZ (N=79) were associated with a higher rate of clinically significant cancer detection when compared to those located in the TZ (N=63) on targeted biopsy (41.8% [33/79] versus 23.8% [15/63], p=0.004). PI-RADSv2 score 5 lesions were more frequent in the TZ compared to score 4 lesions (60.7% [34/56] versus 25.4% [15/59], p=0.003). Therefore, we performed subgroup analyses comparing clinically significant cancer detection rates among PI-RADSv2 score 4 and 5 lesions by location in the PZ and TZ. When controlling for zone location, there was no difference in clinically significant cancer detection rates in PI-RADSv2 score 4 versus 5 lesions in the PZ (Score 4; 34.3% [12/35] cancers versus Score 5; 50.0% [12/24], p=0.34) or the TZ (Score 4; 19.0% [4/21] versus Score 5; 25.7% [9/35], p=0.56).

Mean lesion size was 16 ± 8 (Range 5 to 38) mm with no difference in detection of clinically significant cancers by lesion size (p=0.84). There was a significantly higher clinically significant cancer detection rate among targeted biopsies performed by the core Radiologists
comprising the prostate biopsy service (36.2% [47/130]) compared to the other Radiologists who performed targeted biopsies during the study period (8.3% [1/12]), p=0.05. The mean number of core biopsies per lesion was 2 ± 2 (range 1 to 10) with no association between cancer detection and number of biopsies (p=0.28). TRUS visible tumors (47.9%, 68/142) were associated with a higher yield of clinically significant cancer detection (42.6% [29/68] clinically significant cancer detection among TRUS visible lesions versus 25.7% [19/74] clinically significant cancer detection among TRUS non-visible lesions, p=0.033) and were associated a higher PI-RADSv2 score (p=0.035) and higher Gleason scores (p=0.042) but TRUS visibility was not associated with lesion location (p=0.78) or size (p=0.27).

Follow-up of the 70 men with PI-RADSv2 score 4 or 5 lesions on mp-MRI with biopsies negative for clinically significant cancers was available in only a limited number of patients due to the relatively short period of time between analysis of results and time of biopsy. Two-year follow-up was available in 17 patients, 1-year follow-up was available in 27 patients and in the remaining 26 patients follow-up was less than 1 year in duration. For those patients with at least 1-year of follow-up: 20.5% (9/44) were subsequently diagnosed with Gleason score 3+4=7 PCa (1 on repeat biopsy and 8 at RP), 4.5% (2/4) were diagnosed with Gleason score 4+3=7 PCa (both on repeat biopsy), 4.5% (2/44) were diagnosed with large volume Gleason score 3+3 after RP, 4.5% (2/44) were treated with radiotherapy, 59.0% (26/44) were placed on, or remained on, active surveillance, and 6.8% (3/44) patients were discharged from care. In patients with less than 1 year of follow-up, no repeat biopsies or other interventions had been initiated at time of data analysis.

**Discussion**

This study reports the clinically significant prostate cancer detection rates in mp-MRI-TRUS guided cognitive fusion biopsies stratified by PI-RADS version 2 assessment categories. We observed a low rate of clinically significant cancer detection in biopsies performed for PI-RADSv2 assessment category 4 (clinically significant cancer likely to be present) and 5 (clinically significant cancer very likely to be present) particularly for lesions located in the TZ. Cancer detection was highest in PZ lesions, those which were TRUS visible and when biopsies were performed by the most experienced operators but did not depend on number of core biopsies performed or lesion size.

The clinically significant cancer detection rates in our study are generally lower than what has been reported for PI-RADSv2 assessment category 4 and particularly for category 5 lesions in previous work using either in-bore or automated MRI-TRUS fusion techniques which range from 30 to 78% for PI-RADSv2 score 4 and 78 to 100% for PI-RADSv2 score 5 lesions. Moreover, although follow-up was only available for patients with a negative targeted biopsy in a minority of men in our cohort, the majority of those who went on to repeat biopsy or definitive treatment had clinically significant cancers diagnosed suggesting that the discord between MRI interpretation and cognitive biopsy result is due to sampling errors at biopsy and
not errors in MRI interpretation. In a previous study by Costa et al., repeated MRI-TRUS automated fusion biopsy in Likert score 5 prostate lesions after an initial round of negative targeted biopsies identified clinically significant cancers in approximately 40% more patients 23.

A prior meta-analysis performed by Wegelin et al. (albeit published before reporting of MRI using PI-RADSv2) demonstrated that cognitive fusion is inferior to both in-bore and automated MRI-TRUS fusion systems 10. Our results which are the first to report Canadian data for clinically significant prostate cancer detection in mp-MRI detected targets using cognitive fusion TRUS-guided biopsies where targets are stratified by the PI-RADS version 2 system confirms results from prior studies which show that cognitive fusion is inferior to in-bore MRI guided or MRI-TRUS automated fusion guided systems for targeted prostate biopsies. Limitations in detection of clinically significant cancer among PI-RADSv2 score 4 and 5 lesions with cognitive fusion biopsy are important and must be appreciated by physicians treating prostate cancer and acknowledged by health-care facilities and funding organizations.

We demonstrated that clinically significant cancer detection rates were particularly low in TZ compared to PZ lesions. This is also an expected observation since TZ lesions are located anteriorly which is further away from the end-firing TRUS probe and lesion visibility decreases as a function of distance from end-firing endoluminal probes 3. Not surprisingly, TZ lesions were less TRUS visible in our study compared to PZ lesions and TRUS visibility was significantly associated with clinically significant cancer detection. The disproportionately higher number of TZ lesions in our study likely reflects bias due to the population studied (i.e. men with previously negative template biopsy but with persistent clinical suspicion of cancer and those on active surveillance). TRUS visible lesions were more likely to be associated with a higher PI-RADSv2 score and higher Gleason scores which is concordant with what has been reported previously 24. Our overall yield for clinically significant cancers in PI-RADSv2 score 4 or 5 lesions overall is comparable to the study by Lai et al. which evaluated cognitive MRI-TRUS fusion in lesions stratified by PI-RADS version 2 scores 25 and our rates TZ lesions alone is comparable to the study by Murphy et al. which showed clinically significant cancer detection in only one third of anterior targets using cognitive fusion 26.

In our study, there was a significantly higher cancer detection rate in the most experienced compared to the least experienced operators. An often-cited limitation of cognitive fusion biopsy is operator experience and, our results support that targeted biopsies performed using cognitive fusion have a higher yield among the most experienced operators. There was no difference in cancer detection by number of core biopsies performed in our study. Most studies performing targeted biopsy of mp-MRI detected lesions report a mean number of core biopsies per target of two 19-21 which is concordant with our data. Increasing the number of biopsies did not improve cancer detection rates in our study which is an expected outcome since our results indicate that if the lesion is TRUS visible it can be effectively sampled whereas if it is not TRUS visible increasing the number of biopsies does not improve the yield of cancer detection. Though
an increased number of biopsies is not associated with increased risk of infection, it has been associated with increased patient morbidity with associations between increased number of biopsies performed and increased post-procedural bleeding and pain\textsuperscript{27}. Our study showed decreased clinically significant PCa detection rates in TZ compared to PZ lesions and no difference in PCa detection by cognitive fusion in larger compared to smaller tumors which is concordant with the prior results reported by Cool et al. \textsuperscript{11}

Our study has limitations. Our patient population all underwent their targeted biopsies fairly recently (within 3 years of data collection and analysis) which limited our ability to perform meaningful long-term follow-up of patients with PI-RADS score 4 or 5 lesions and negative biopsy. This limitation is expected given that targeted biopsies of MRI detected lesions have only recently become a part of the standard of care management pathways for men treated with active surveillance or with negative template biopsies and persistent suspicion of cancer. In our study, a majority of these men were either enrolled in or remained on active surveillance and clinically significant cancer diagnosis was established only in a minority of these men during follow-up.

In conclusion, we observed a low rate of clinically significant prostate cancer diagnosed on cognitive fusion mp-MRI TRUS guided biopsy in PI-RADS\textsuperscript{v2} category 4 and 5 lesions compared to rates of cancer detection reported in the literature using either in-bore or automated MRI-TRUS fusion systems, particularly in transition zone tumors. Our lower rates of clinically significant cancer detection are comparable to rates described by other authors using cognitive fusion as a method of guiding targeted prostate biopsies. There was improved detection of significant cancers in TRUS visible and peripheral zone lesions and in the most experienced operators but no difference in yield by number of additional core biopsies performed or size of tumor. Our results have several important implications for delivery of patient care in Canada. Firstly, when combined with the other Canadian study by Cool et al. \textsuperscript{11}, it can be concluded that cognitive MRI-TRUS guided targeted biopsies is not as accurate for sampling of MRI detected lesions compared to other targeting systems and yields lower than expected rates of significant cancers when MR imaging findings are stratified by PI-RADS version 2. Canadian health care facilities and government must invest capital for the acquisition of alternative targeting systems (e.g. automated MRI-TRUS fusion systems or in-bore MRI directed guidance systems) to maintain accuracy of prostate cancer diagnosis compared to other developed countries. Secondly, when an MRI detected lesion is not TRUS visible, a physician receiving targeted biopsy results should be aware of the low cancer detection rates from cognitive fusion particularly in TZ lesions. Negative cognitive biopsy results in this setting should not be considered as truth and level of suspicion of cancer based upon MRI results (e.g. PI-RADS score) and other clinical parameters should guide management decisions for patients. Third, increasing the number of core biopsies to sample a target that is not TRUS visible did not improve yield in our study and
we suggest that no more than 2-3 targeted biopsies be performed per target, as is the clinical standard currently, to minimize patient procedure related morbidity.
References


13. Rozenberg R, Thornhill RE, Flood TA, Hakim SW, Lim C, Schieda N. Whole-Tumor Quantitative Apparent Diffusion Coefficient Histogram and Texture Analysis to Predict

Table 1. Prostate cancer detection in 142 lesions in 131 men stratified by PI-RADS version 2 assessment category and location among targeted biopsies performed using cognitive fusion of MRI and TRUS

<table>
<thead>
<tr>
<th>PI-RADSv2 assessment category* 3 (n=27)</th>
<th>PI-RADSv2 assessment category 4 (n=56)</th>
<th>PI-RADSv2 assessment category 5 (n=59)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peripheral zone (n=23)</td>
<td>Peripheral zone (n=35)</td>
<td>Peripheral zone (n=24)</td>
</tr>
<tr>
<td>Transition zone (n=4)</td>
<td>Transition zone (n=21)</td>
<td>Transition zone (n=35)</td>
</tr>
</tbody>
</table>

Any cancer diagnosis (including Gleason score $3+3=6$ or higher)

- 30.4% (7/23) Periphera
- 25.0% (1/4) Transition
- 60.0% (21/35) Periphera
- 66.7% (14/21) Transition
- 70.8% (17/24) Periphera
- 57.1% (20/35) Transition

Clinically significant cancer (Gleason score $3+4=7$)

- 13.0% (3/23) Periphera
- 0 Transition
- 34.3% (12/35) Periphera
- 19.0% (4/21) Transition
- 50.0% (12/24) Periphera
- 25.7% (9/35) Transition

No cancer diagnosed

- 69.6% (16/23) Periphera
- 75.0% (3/4) Transition
- 40.0% (14/35) Periphera
- 33.3% (7/21) Transition
- 29.2% (7/24) Periphera
- 42.9% (15/35) Transition

PI-RADS version 2 assessment categories were assigned by an experienced radiologist where assessment category 3=likelihood of clinically significant cancer is indeterminate or equivocal, 4=likelihood of clinically significant is high, and 5=likelihood of clinically significant cancer is very high. MRI: magnetic resonance imaging; PI-RADS: Prostate Imaging and Data Reporting System; TRUS: transrectal ultrasound.