Detrimental Variations in the Interpretation of Pre-operative Prostate MRI Reports Reinforces the Need for a Standard Report Structure

Poster No.: C-1301
Congress: ECR 2014
Type: Scientific Exhibit
Keywords: Genital / Reproductive system male, MR, Observer performance, Decision analysis, Patterns of Care
DOI: 10.1594/ecr2014/C-1301

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method is strictly prohibited.

You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.

www.myESR.org
Aims and objectives

The purpose of this study is to examine, in the absence of a structured reporting system, the consistency with which clinicians draw conclusions from reading magnetic resonance imaging (MRI) reports on prostate cancer staging in order to make appropriate therapeutic pathway choices. Magnetic resonance imaging has demonstrated clinical value in determining the stage of prostate cancer for surgical, and, potentially, for radiation therapy hypo-fractionation planning [1]. The current processes for prostate tumor screening include digital rectal examination (DRE) and obtaining prostate specific antigen (PSA). If PSA results exceed 3-4 ng/mL, trans-rectal ultrasound image guided biopsies are usually conducted to assess presence and gland locale of cancer. In general, PSA has a low specificity, which results in a high proportion of false positives. Multi-parameter MRI has proven to have a higher specificity in determining presence of a tumor as well as extent of the disease. More specifically, MRI has been reported to be an important tool for determining cancer extra-capsular extension (ECE) as well as involvement of the neurovascular bundle (NVB) [2,3]. Extra-capsular extension is often accompanied by a Gleason pathology score greater than 7 or an increasing PSA level more than 20 ng/ml [4].

Patients diagnosed with cancers not confined within the prostatic capsule are often recommended for non-surgical treatment options such as radiation oncology. On the other hand, patients that have localized cancer are eligible for radical prostatectomy or, if small, active surveillance. Depending on how certain it is that the cancer has spread to the nerves, the NVB may or may not be surgically excised at the time of prostatectomy. Studies have shown that patients with nerve sparing surgery have lower adverse events such as incontinence and erectile dysfunction post-surgery [5]. It is for this reason that MRI assessment of ECE and/or involvement of NVB be uniformly assessed with the same assurance by all MR imaging readers. Patients are generally divided into three major categories depending on the histological Gleason score and PSA levels [6]. Low risk patients are those with a Gleason total score less than or equal to 6, and PSA levels less than 10 ng/mL [6]. Patients with intermediate risk tend to have a PSA between 10 and 20 ng/mL and a total Gleason score of 7 [6]. High risk patients are commonly defined by a PSA greater than 20 and a Gleason score of 8-10 [6].

Currently, a standard method for MRI-based reporting of stage has been promoted by the European Society of Urogenital Radiology known as the PI-RADS scoring system [7]. Breast radiologists implement a similar system known as BI-RADS to assess the likelihood of metastatic disease. The scoring methods for PI-RADS are based on a scale of 1-5. For example, a score value of 1 defines an absence of cancer. As the ratings increase score 5 defines a high likelihood of cancer being present [6]. This reporting system has been reported as useful for accurately predicting histology in pre-surgical and active surveillance patients [7-9].
In an effort to create a standardized dictionary of semantic terms and a useful indexing scheme for data archives of images and imaging reports, RadLex™ has been developed to assist radiologists worldwide to communicate by uniform definitions [11]. Specifically, RadLex Playbook™ is a centralized resource where definitions can be found relating to "devices, imaging exams, and procedure steps", enabling universal communication between physicians to be as clear as possible [11]. Radlex also works towards creating a standard reporting system so that when MRI cases are diagnosed for research and educational purposes, a clinical report with uniform meaning is provided to all readers [11].

The purpose of this study was to evaluate differences in interpretation of prostate MRI reports between radiologists with particular regard to the likelihood of extra-capsular extension and involvement of neurovascular bundles. Information on these particular structures has a key influence on treatment decisions and surgical planning. Specifically, eligibility for nerve sparing surgery and overall MRI staging was included as an essential element of each report.

**Methods and materials**

Following IRB approval, from January 2012-July 2013, five radiologists reviewed 54 MRI reports of consecutive patients who underwent pre-surgical prostate-MRI using a high resolution contrast-enhanced MRI protocol (HR-CEMRI).

**HR-CEMRI preparations**

Prior to HR-CEMRI acquisition, standard anesthesia protocol was followed and included local anesthesia of the anal region with topical lidocaine gel application was performed just before the HR-CEMRI examinations (Xylocaine 2% Gel, Astra Zeneca, Wedel, Germany). To reduce bowel peristalsis, 0.5 mg Glucagon (Glucagen, Novo Nordisc, Bagsvaerd, Denmark) was administered IV just before the HR-CEMRI examinations and 0.5 mg during the examinations (added to saline syringe of the automated injection system).

**HR-CEMRI Protocol**

Prostate HR-CEMRI examinations were performed on a 1.5T scanner with a pelvic phased-array surface coil combined with a disposable endorectal prostate coil (eCoil, MEDRAD, Inc., Pittsburgh PA). The endorectal coil was connected to the pelvic phased-array surface coil, and combined signal images were obtained. All images were analytically corrected for the reception profile of the endorectal and pelvic phased-array coils. Sagittal and transverse half-Fourier single-shot turbo spin-echo (TSE) sequences were first obtained to check coil position. High-resolution T1- and T2-weighted images
were obtained in the transverse plane with a 16-cm field of view (FOV), matrix with 256 frequency-encoding steps, and 192 phase-encoding steps yielding in-plane spatial resolution of 0.63 x 0.93 mm, phase direction right-left, 100% phase over sampling. The axial T2-weighted sequence was a dual echo TSE sequence acquired from below the apex of the prostate to above the seminal vesicles with the following parameters: repetition time msec/first effective echo time msec/ second effective echo time msec 4000/83/165, echo train length of 8.3, slice thickness (ST) 3 mm, no intersection gap, 28 slices, three signals averaged (acquisition time: 10min,48s). T1-weighted images were acquired before and after contrast administration.

Report Interpretation

Based on individual interpretation of each MRI report the probability of ECE, seminal vesicle (SV) and eligibility for nerve sparing surgery (NSS) was recorded by each reader. Blinded readers were provided two sets of the same 54 MRI reports shuffled in order and completed 1-week apart to assess for intra-reader variability. At the interpretation of each report, the radiologists were asked to assess the probability of ECE and SV involvement on a scale from 1-5. A Likert scale rating for presence of tumor involvement of ECE and SV was: 1) Not present 2) Probably not present 3) Possibly present 4) Probably present and 5) Definitely present as we have used in a prior study. Lastly, the interpreting radiologists were asked to score their recommendation for the patient’s ability to undergo NSS on each side based on the information provided in the report. Eligibility for undergoing NSS was assessed on a scale from 1 to 3. The point scale rating for this study was: 1) not recommended 2) indeterminate based on MRI report and 3) recommended.

Statistical Analysis

Descriptive analyses were performed to obtain patient, tumor and treatment characteristics. The inter-rater agreement between five interpreting radiologists was assessed by calculating Intra-class correlations (ICCs) for each structure. Intra-class correlations were calculated for each of the following: the probability of extra-capsular extension, involvement of the seminal vesicles, and the eligibility of the patient to be considered for left nerve sparing surgery and right nerve sparing surgery. A two-way random effects model (absolute agreement) was used to calculate ICCs. The correlations were calculated for each round of reports. A p-value of 0.05 or less was considered to indicate a statistically significant difference. All statistical analyses were performed using IBM SPSS Statistics software, (version 20.0; IBM Corp) and SAS 9.1 system (SAS Institute, Cary, NC).

Results
**Patient Demographics**

All 54 patients included in this study underwent total prostatectomy. Forty (74.7%) patients were diagnosed with pathologic (p) stage T2c, followed by seven (12.7%) patients with pT3a, five (9.1%) patients with pT2a, and 2 (3.6%) patients with pT3b cancer. Patient ages ranged from 40 to 78 years of age (mean 61 year old, standard deviation 7.6 years). Figures 1-3 are examples of patients included in this set of 54 patients undergoing pre-operative HR-CEMRI assessment. Excerpts from the dictated reports are included with the figures and are examples of the reports interpreted by the 5 radiologists.

**HR-CEMRI Report Interpretation**

ICCs were statistically significant for each structure assessed and an improvement in inter rater agreement was noted on round 2 compared to round 1, Table 1. ICCs for ECE involvement were 64% and 72% for round 1 and 2, respectively. Similar trends reflecting an improvement in the ICCs between rounds 1 and 2 were also seen for the SV whereby ICCs for round 1 was 6% compared to 31% for round-2. This trend was also observed for the assessment of the patients to undergo nerve sparing surgery. The ICCs for left-sided NSS was 61% on round 1 compared to 75% for round 2. ICC for right-sided NSS demonstrated mild improvement between the two rounds at 62% on round 1 compared to 66% on round-2.

**Images for this section:**
**Fig. 1:** Figure 1a and 1b: Patient 1 Example patient undergoing pre-operative assessment with HR-CEMRI. The dictated report describes a large region of T2 hypointensity and suspicious enhancement involving predominantly the left peripheral zone (yellow arrow) which extends to and involves the seminal vesicles (red arrow).

![MRI Image](image_url)

**Fig. 2:** Figure 2: Patient 2 Example patient undergoing pre-operative HR CE-MRI assessment. The dictated report describes a 1.6cm prostate cancer in the left peripheral zone with imaging findings highly suggestive for beginning extracapsular disease extending towards the neurovascular bundle (radial extension of 1mm beyond the capsule)(yellow arrow). The right neurovascular bundle and seminal vesicles are uninvolved. This patient also has co-existent benign prostatic hypertrophy (BPH).
Fig. 3: Figure 3: Patient 3 Example patient undergoing pre-operative HR-CEMRI assessment. The dictated report describes a 9mm suspicious nodule in the right posterior lateral peripheral zone at 7 o'clock (yellow arrow), at the prostate base, with capsular infiltration and possible beginning ECE. There is maximal 1-2mm radial extension-towards the right neurovascular bundle. The left neurovascular bundle uninvolved. No evidence of seminal vesicles infiltration.
| Round 1: Likelihood of Extracapsular Extension (1-5) | 55 | 0.64 (0.52 – 0.74) | <0.01 |
| Round 2: Likelihood of Extracapsular Extension (1-5) | 55 | 0.72 (0.60 – 0.81) | <0.01 |
| Round 1: Likelihood of Seminal Vesicle Involvement (1-3) | 55 | 0.06 (-0.01 – 0.15) | 0.05 |
| Round 2: Likelihood of Seminal Vesicle Involvement (1-3) | 55 | 0.31 (0.19 – 0.45) | <0.01 |
| Round 1: Nerve Sparing Surgery on L (1-3) | 55 | 0.61 (0.50 – 0.72) | <0.01 |
| Round 2: Nerve Sparing Surgery on L (1-3) | 55 | 0.75 (0.66 – 0.83) | <0.01 |
| Round 1: Nerve Sparing Surgery on R (1-3) | 55 | 0.62 (0.50 – 0.73) | <0.01 |
| Round 2: Nerve Sparing Surgery on R (1-3) | 55 | 0.66 (0.56 – 0.76) | <0.01 |

**Table 1**: Table 1: Depicts the intra-class correlations (ICCs) for each of the structures assessed by the five interpreting radiologists. The confidence intervals (CI) and P-values are listed for each structure assessed. ECE = extra-capsular extension, SV = seminal vesicle, L NSS = left nerve sparing surgery, R NSS = right nerve sparing surgery.
Conclusion

This study indicates that there is great variability in the understanding of the clinical implications of the MRI descriptions between radiologists when they read free-text MRI reports. Due to a lack of standardized reporting, radiologists were asked to infer stage implications based on semantics, wording, and previous knowledge. One major observation revealed in this study was that there was a slight improvement in intra-class correlations between the first round and second round of interpretations. Intra-class correlations (ICC) for ECE involvement increased from 64% (round-1) to 72% (round-2). ICCs for SV involvement rose from 6% (round-1) to 31% (round-2). ICC for left-sided NSS was 61% (round-1) to 75% (round 2). ICC for right-sided NSS was 62% (round-1) to 66% (round-2). This suggests that even though both rounds of reports were administered a week apart, there was some influence of familiarity associated with the wording between the two groups. The greatest differences in ICC were seen in the evaluation of SV involvement. In general recommendation on nerve sparing surgery showed less variability and readings were quite similar for both sides of the body. ECE involvement showed greater correlation in both rounds. This suggests that the free-text MRI reports are much more successful at addressing ECE involvement and indications for NSS. Semantic inferences about involvement of the SV were much less clear and therefore is an important element of reports that needs to be improved. These results suggest that readers are able to extract almost the same information when reading about ECE and NSS, but are much more varied on implications SV involvement. This means that the standardized reporting structure should include some of the terms that are used to distinguish different benchmarks for ECE and NSS, but needs to develop a new framework for expressing SV involvement.

The lack of implementation of a standardized reporting system can be a serious obstacle to consistent patient care. Our study shows the interpretation of routinely free-text prostate MRI radiology reports result in a wide range of reader variability regarding intended exam conclusions about ECE and tumor involvement of NVB. These differences in interpretation can eventually lead to an incorrect choice of treatment options. Prior studies have explored the concept of a standardized MRI reporting system. Rosenkrantz et al. found that using the PI-RADS system proved to produce greater interreader reproducibility in the peripheral zone as compared to the transition zone and that more experienced readers had less variance than less experienced readers [10]. This suggests that though the PI-RADS system has been found to be successful to some extent in conveying a certain message, it still lacks clarity in expressing a message successfully in every part of the organ. It is vital that any system created be adept for clarity in all regions and that radiologists of all experiences be able to extract the same information from a report. In another study conducted by Rosenkrantz AB, two standardized systems, PI-RADS and Likert scale, were compared and it was found that both systems were found to be successful but the transition zone was found to be more clearly described through the Likert scale [10]. This again emphasizes the importance of making sure that
a standardized system be able to communicate the localization of, or lack of, disease accurately in all areas of the prostate. Overall, it is essential that specificities be made for each region of the organ so that accurate, understandable descriptions can be formed and accordingly reports can be interpreted correctly.

An imaging report is of utmost importance to patient care for it not only should provide an understandable record, but also acts as a communication media between clinicians. It is vital that the report not only convey a consistently understood message between physicians, but also to patients. It would also enable a reliable index for subsequent health care outcome data research. By developing a uniform dictionary of radiology terms physicians from all over the globe have a greater chance of fully understanding each other’s work, as well as having the means to form more meaningful collaborations. In addition to benefits from a research point of view, a standardized template has clinical benefits as well. In a study conducted, it was found that the ideas that were mentioned on a standardized template are also found in the free-text clinical reports [12]. This suggests that developed templates do take into account full coverage of the material that is found within free-text reports, but offer the benefit of identifying main points and to access for standards that need to be met globally.

This study proves the importance of developing a standardized structured reporting system with universal qualitative descriptors of intra-and extra-prostatic disease staging linked with defined implications for treatment-specific options. This will help avoid misinterpretation leading to unfavorable treatment pathways with respect to active-surveillance, NSS, and prostatectomy versus radiation-therapy. Because there is no standardized reporting system and qualitative descriptors are not clearly defined or universally used, individual interpretations of MRI reports may lead to misinterpretation of tumor-stage. As a result, inappropriate treatment options could be advised to patient, and optimal outcome may be imperiled.

The need for standardized reporting schema is vital to communicate a consistent framework for describing patients of similar disease status. It also ensures that physicians from various disciplines are able to comprehend the same report consistently and ultimately choose the most appropriate treatment pathway for the patient.

Personal information

References

REFERENCES


