Long-distance longitudinal prostate MRI quality assurance: from startup to 12 months

Nicole E. Curci,[1] Patrick Gartland,[2] Prasad R. Shankar,[1] Jeffrey S. Montgomery,[3]
David C. Miller,[3] Arvin K. George,[3] Matthew S. Davenport[1,3,4]

[1]Department of Radiology, Michigan Medicine, Ann Arbor, MI, USA

[2]Grand Traverse Radiologists PC, Munson Medical Center, Traverse City, MI, USA

[3]Department of Urology, Michigan Medicine, Ann Arbor, MI, USA

[4]Department of Radiology, Michigan Radiology Quality Collaborative, 1500 E. Medical Center Dr. B2-A209P, Ann Arbor,
MI 48108, USA

Abstract

Purpose: To evaluate a 12-month long-distance prostate MRI quality assurance (QA) program.
Methods: The need for IRB approval was waived for this prospective longitudinal QA effort. One academic institution experienced with prostate MRI [~ 1000 examinations/year (Site 2)] partnered with a private institution 240 miles away that was starting a new prostate MRI program (Site 1). Site 1 performed all examinations ($N = 249$). Four radiologists at Site 1 created finalized reports, then sent images and reports to Site 2 for review on a rolling basis. One radiologist at Site 2 reviewed findings and exam quality and discussed results by phone (~ 2–10 minutes/MRI). In months 1–6 all examinations were reviewed. In months 7–12 only PI-RADS ≤ 2 and 'difficult' cases were reviewed. Repeatability was assessed with intra-class correlation (ICC). 'Clinically significant cancer' was Gleason ≥ 7.
Results: Image quality significantly ($p < 0.001$) improved after the first three months. Inter-rater agreement also improved in months 3–4 [ICC: 0.849 (95% CI 0.744–0.913)] and 5–6 [ICC: 0.768 (95% CI 0.619–0.864)] compared to months 1–2 [ICC: 0.621 (95% CI 0.436–0.756)]. PI-RADS ≤ 2 examinations were reclassified PI-RADS ≥ 3 in 19% (30/162); of these, 23 had post-MRI histology and 57% (13/23) had clinically significant cancer (5.2% of 249). False-negative examinations [$N = 18$ (PI-RADS ≤ 2 and Gleason ≥ 7)] were more common at Site 1 during months 1–6 [9% (14/160) vs. 4% (4/89)]. Positive predictive values for PI-RADS ≥ 3 were similar.
Conclusion: Remote quality assurance of prostate MRI is feasible and useful, enabling new programs to gain durable skills with minimal risk to patients.

Key words: Quality—Prostate MRI—Remote quality assurance—Collaboration—Long-distance

Multiparametric prostate MRI has emerged as an accurate and reliable method of detecting and risk-stratifying primary prostate cancer [1–3], both serving as a tool for diagnosis and as a means of facilitating targeted biopsy through either cognitive fusion or commercially available fusion platforms coupled with transrectal ultrasound [3]. The Prostate Imaging Reporting and Data System (PI-RADS), now in version 2 [1, 2], has been developed to help educate providers, ensure quality of image acquisition, and standardize reporting methods. However, performance and interpretation are not simple [4–12], and new programs seeking to implement this new technology may struggle to integrate it. This can be particularly true in smaller practices staffed by radiologists who may not have received dedicated instruction in this technique during their training. It has been shown that even in expert hands, interpretation is challenging [4], and the technique is not necessarily 'plug-and-play.'

Given the rapid rise in prostate MRI utilization across the world [12–14], developing ways to train new and existing untrained radiologists, and doing so efficiently without detriment to patient care, is of utmost

importance. Although weekend courses and brief hands-on workshops exist, it is unclear that these result in durable retention of practical knowledge. Such courses are unable to document longitudinal quality of care, and the imaging to be reviewed is hand selected by the course directors to be provided to the participants—not acquired in real time. In a real-world scenario, imaging protocols must be designed, tested, and optimized. Learning on the job can negatively affect patients if non-proctored education is occurring simultaneously with patient care. Additionally, if the learning process is not tracked and monitored, the care may never reach an optimal state [15].

Direct in-person mentored feedback akin to the processes inherent to radiology residency and fellowship programs has been the traditional model of teaching new complex diagnostic radiology skills [8], but this may not be practical for busy or remote practices. Ideally, practices adopting a new technique like prostate MRI would have consistent monitored oversight by trained providers to prevent harm to patients, teach necessary skills, and create lifelong portable expertise. In light of these goals and challenges, we partnered with a remote private practice institution in a long-distance 12-month longitudinal quality assurance training program designed to optimize prostate MR imaging technique and teach quality prostate MRI interpretation skills. The purpose of this report is to demonstrate the impact and utility of that program.

Methods

This prospective longitudinal quality assurance effort was considered to be exempt from oversight by the host institutional review board (IRB).

Quality assurance process

One academic institution experienced with prostate MRI [$n \sim 1000$ examinations/year (i.e., supervising institution)] partnered with a private practice institution 240 miles away that was starting a new prostate MRI program (i.e., in-training institution). The relationship was mutually desired and contractual (12 months). The in-training site performed all examinations ($n = 249$), and paid the supervising site roughly the equivalent of the sum of the professional fees for each examination reviewed. Four radiologists at the in-training institution created finalized reports without immediate oversight, then sent images and reports to the supervising institution for review on a rolling basis. Each of the four radiologists at the in-training institution was board-certified and had experience reading clinical body MRI; one had completed a fellowship in general MRI. The finalized reports contained disclaimer language indicating that the images would be reviewed for quality assurance and that an addendum would be made once that review

was complete. One radiologist with expertise in prostate MRI (> 2000 prostate MRIs reviewed) at the supervising institution reviewed the imaging, original MRI report, and exam quality as the examinations were sent and discussed those results with the in-training radiologist(s) by phone. The time the supervising radiologist spent to review and interpret each examination was similar to the time required to review a prostate MRI clinically (i.e., 10–20 min/MRI). Any clinically significant changes arising from that review were then appended to the original reports and discussed with the referring providers by the in-training radiologists.

For all examinations discussed, MR technical parameters, image quality, image optimization, lesion detection and location, and lesion characteristics were reviewed. These phone calls lasted roughly 2–10 min per MRI discussed, with longer conversations earlier in the quality assurance process. Phone calls occurred approximately weekly or biweekly. In months 1–6 all examinations were reviewed ($n = 160$). In months 7–12 only PI-RADS ≤ 2 and 'difficult' cases were reviewed ($n = 89$). 'Difficult' examinations were those the in-training radiologists determined they had subjective difficulty interpreting.

The goal of the quality assurance process was to develop durable skills in the in-training radiologists and technologists that would enable reliable and excellent image quality and interpretations. None of the radiologists at the in-training institution had previously undergone formal training in multiparametric prostate MRI interpretation.

MRI examinations and interpretations

All multiparametric prostate MRI examinations were performed according to PI-RADS v2 guidelines [1] on a 3.0-Tesla magnet (Skyra, Siemens Healthcare, Erlangen, Germany) without an endorectal coil. The in-training technologists and radiologists reviewed the PI-RADS v2 guidelines [1] prior to initiating the quality assurance program. As the program progressed, iterative improvements in scan technique were implemented. The following sequences were acquired: narrow field-of-view (prostate) multiplanar (sagittal, axial, coronal) 2D T2-weighted turbo spin echo, narrow field-of-view (prostate) axial 3D T2-weighted turbo spin echo, narrow field-of-view (prostate) axial diffusion-weighted imaging with apparent diffusion coefficient map (b-values: 0, 400, 800, 1600 s/mm^2), narrow field-of-view (prostate) axial dynamic contrast-enhanced imaging with a temporal resolution of 6–7 s, large field-of-view (whole pelvis) axial T1-weighted dual-echo gradient-recalled echo with Dixon reconstructions pre- and post-contrast. Dynamic contrast-enhanced imaging was obtained following the bolus administration (2 mL/s) of 0.1 mmol/kg gadobutrol (Gadavist, Bayer HealthCare Pharmaceuticals,

Whippany NJ). The exact technical parameters of each sequence varied throughout the quality assurance effort as image quality was optimized.

All examinations were interpreted at the in-training and supervising sites using the same structured reporting template (Appendix A) created in Powerscribe 360 (Nuance Communications Inc, North Sydney, Australia). Full reports were created at both sites. The template included structured elements related to the prostate (i.e., calculated gland volume [length * width * height * 0.52 mL], subjective volume of benign prostatic hyperplasia, subjective volume of median lobe enlargement), periprostatic anatomy (i.e., membranous urethra length on coronal imaging [16], lymph nodes, bones), and prostate lesion characteristics (i.e., number, location, size, imaging features by sequence, morphology, length of capsular contact [17], local extent, PI-RADS v2 score [1]). Identical structured reporting allowed direct comparison between the in-training and supervising institution reports.

Data analysis

Descriptive statistics were calculated. Inter-rater repeatability was assessed per quarter and every two months with intra-class correlation coefficients (ICC) for continuous data (e.g., maximum PI-RADS v2 score) and with Kappa statistics for categorical data (e.g., subjective volume of benign prostatic hyperplasia). Kappa results (slight: 0.01–0.20, fair: 0.21–0.40, moderate: 0.41–0.60, substantial: 0.61–0.80, almost perfect: 0.81–0.99 [18]) and ICC results (poor: < 0.40, fair: 0.40–0.59, good: 0.60–0.74, excellent: 0.75–1.00 [19]) were stratified qualitatively. 'Clinically significant prostate cancer' was defined as Gleason ≥ 7. The reference standard for Gleason ≥ 7 prostate cancer was any histology (sextant biopsy, targeted biopsy, radical prostatectomy) obtained at either the in-training site or supervising site within 6 months after MRI. Positive and negative predictive values were calculated using only those with available histology.

Results

The details of the study population are shown in Table 1. Median patient age was 68 years (IQR: 63–72), and median PSA was 6.8 ng/mL (IQR: 4.9–9.3 ng/mL). Most MRIs were performed in patients with an elevated PSA and negative prior biopsy(-ies) [42% (105/249)] or in those with low-risk prostate cancer being considered for active surveillance [36% (89/249)] (Table 1, Fig. 1). Im-

Table 1. Study population details

Characteristic	All patients	In-Training institution	Supervising institution
MRI examinations (N)	249		
Indication for MRI			
Elevated PSA, negative biopsy (-ies) (%)	42% (105)		
Low-risk cancer, active surveillance (%)	36% (89)		
Pre-operative planning (%)	14% (36)		
Prior local therapy, suspected local recurrence (%)	3% (7)		
Other (%)	5% (12)		
Median age, years (IQR)	68 (63–72)		
Median pre-MRI PSA (ng/mL, IQR)	6.8 (4.9–9.3)		
Histology within 6 months after MRI			
None (%)	63% (158)	–	–
Standard biopsy (%)	31% (78)	–	–
Targeted biopsy (%)	23% (58)	–	–
Standard and targeted biopsy (%)	22% (56)	–	–
Radical prostatectomy (%)	8% (19)	–	–
Median prostate volume (mL, IQR)		53 (39–73)	54 (38–80)
Median membranous urethra length (mm, IQR)		13 (10–15)	14 (11–16)
Number of radiologists		4	1
Practice type		Private practice	Academic
Maximum PI-RADS score (%)			
1 or 2 (%)		65% (162)	57% (143)
3 (%)		6% (15)	9% (22)
4 (%)		15% (38)	14% (36)
5 (%)		11% (27)	16% (41)
Not assigned (prior prostatectomy) (%)		3% (7)	3% (7)
Post-MRI Histology (within 6 months)			
None (%)	63% (158)		
Benign (%)	8% (19)		
Gleason 3 + 3 = 6 (%)	11% (27)		
Gleason ≥ 7 (%)	18% (45)		
Gleason 3 + 4 = 7 (%)	12% (29)		
Gleason 4 + 3 = 7 (%)	4% (10)		
Gleason 8–10 (%)	2% (6)		

Numbers in parentheses are *N* or inter-quartile ranges (IQR)
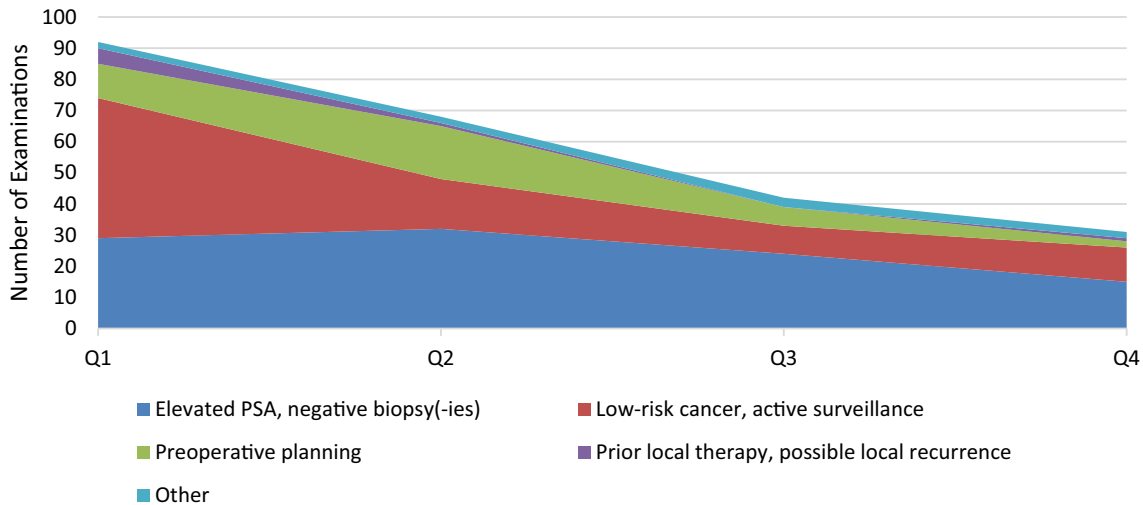*PSA* prostate-specific antigen

Fig. 1. Indications for use of prostate MRI by quarter. *PSA* prostate-specific antigen.
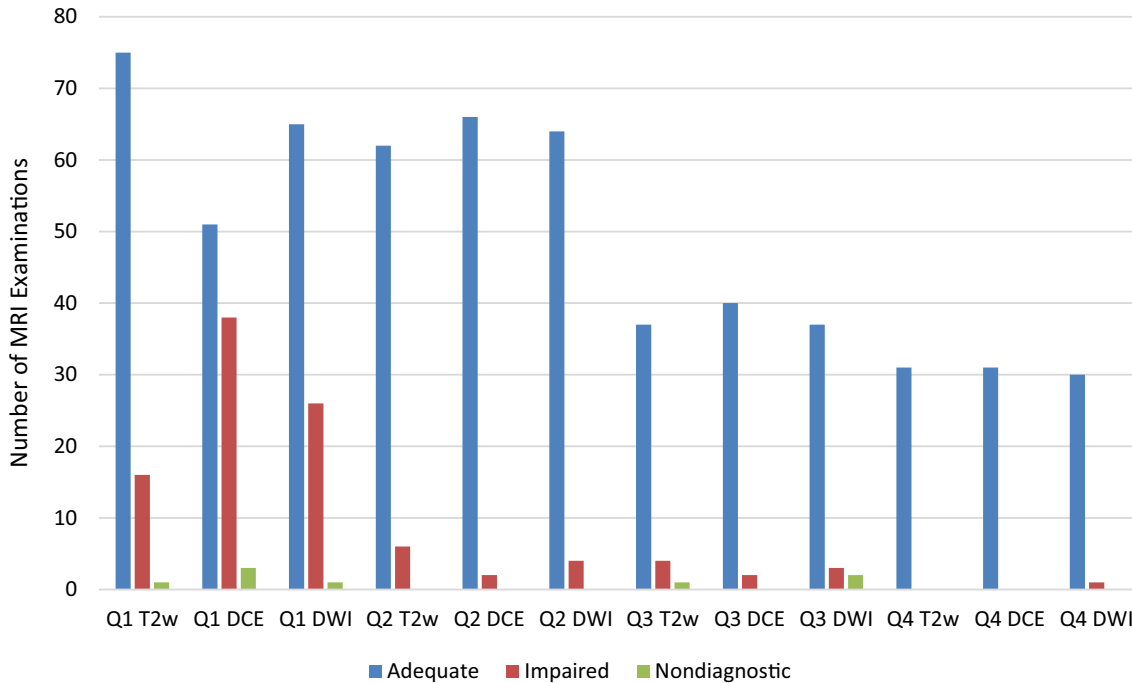

Fig. 2. Diagnostic quality of prostate MRI examinations by quarter and multiparametric sequence. *T2w* high-resolution T2-weighted. *DCE* dynamic contrast-enhanced. *DWI* diffusion-weighted imaging.

age quality was significantly ($p < 0.001$) more likely to be impaired or non-diagnostic in the first three months (Fig. 2). Seven examinations [3% (7/249); 6 in months 1–6, 1 in months 7–12] were not assigned a PI-RADS v2 score because they were performed following local therapy (Table 1).

Three in-training radiologists participated in quarters 1 and 2. In quarters 3 and 4, when only 'negative' and 'difficult' examinations were reviewed, a fourth in-training radiologist was added due to a need at the in-training

site to create redundant skill sets (Table 2). Inter-rater agreement for maximum PI-RADS score was lower in months 1–2 [ICC: 0.621 (95% CI 0.436–0.756)] than in months 3–4 [ICC: 0.849 (95% CI 0.744–0.913)] and 5–6 [ICC: 0.768 (95% CI 0.619–0.864)] (Fig. 3). Agreement for maximum PI-RADS score was highest in quarter 2 (ICC: 0.82 {excellent}, Table 2) and declined in quarters 3 and 4 when the types of cases being reviewed were restricted (i.e., a subgroup not reflective of all PI-RADS scores or examinations) (Table 2). The types of cases

Table 2. Examinations interpreted by reader and agreement statistics by quarter using single-rater mixed-model intra-class correlation statistics (1) for continuous measures and Kappa statistics (2) for categorical measures

Characteristic	Quarter 1 ($N = 92$)	Quarter 2 ($N = 68$)	Quarter 3 ($N = 42$)	Quarter 4 ($N = 47$)
Number of examinations interpreted				
In-training radiologist 1 (%)	40% (37)	50% (34)	29% (12)	28% (13)
In-training radiologist 2 (%)	42% (39)	37% (25)	38% (16)	34% (16)
In-training radiologist 3 (%)	17% (16)	13% (9)	7% (3)	15% (7)
In-training radiologist 4 (%)	0% (0)	0% (0)	26% (11)	23% (11)
Supervising radiologist 1 (%)	100% (92)	100% (68)	100% (42)	100% (47)
Non-neoplastic findings				
[1]Calculated prostate volume	0.95 (0.93–0.97)	0.92 (0.88–0.95)	0.95 (0.91–0.97)	0.95 (0.92–0.97)
[2]Subjective volume of BPH	0.74 (0.67–0.81)	0.79 (0.72–0.86)	0.80 (0.72–0.88)	0.85 (0.77–0.92)
[2]Degree of median lobe enlargement	0.60 (0.53–0.67)	0.57 (0.51–0.63)	0.64 (0.55–0.73)	0.61 (0.53–0.69)
[1]Membranous urethra length	0.29 (0.06–0.48)	0.32 (0.08–0.52)	0.25 (0.00–0.52)	0.42 (0.14–0.64)
Neoplastic findings				
[1]Maximum PI-RADS score	0.68 (0.55–0.78)	0.82 (0.72–0.88)	0.64 (0.43–0.79)	0.39 (0.11–0.61)
[2]Extracapsular extension	0.42 (0.07–0.77)	0.34 (0.08–0.61)	0.55 (0.15–0.94)	0.64 (0.19–0.99)
[2]Seminal vesicle invasion	0.31 (0.00–0.80)	0.64 (0.27–0.99)	–	–

Agreement statistics compare the interpretations of the in-training radiologists to those of the supervising radiologist. All MRI examinations were reviewed in quarters 1 and 2. Only 'difficult' and 'negative' MRI examinations were reviewed in quarters 3 ($N = 9$ 'difficult', $N = 33$ 'negative') and 4 ($N = 4$ 'difficult', $N = 43$ 'negative')
BPH benign prostatic hyperplasia
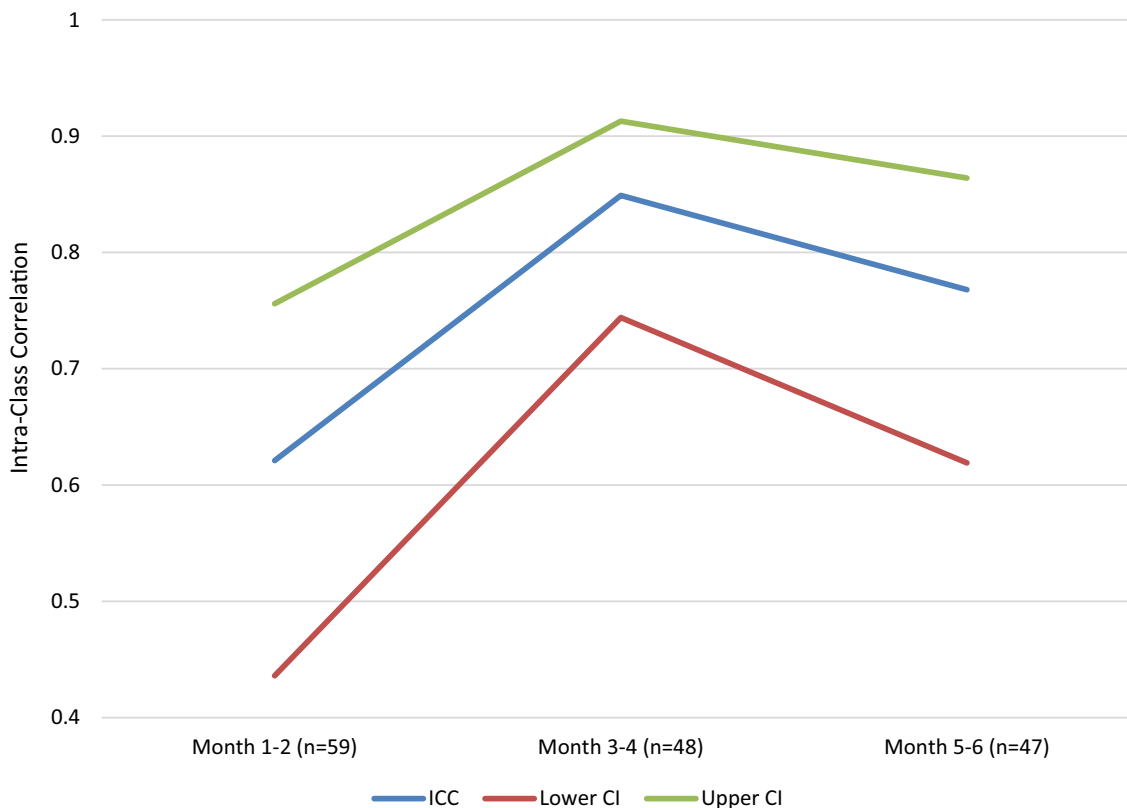– : Data insufficient for analysis (i.e., low event rate)

Fig. 3. Bimonthly inter-rater agreement of maximum assigned PI-RADS score within the first two quarters. Agreement during quarters 3 [ICC: 0.64 (0.43–0.79)] and 4 [ICC: 0.39 [0.11–0.61]] is biased because it was intentionally constrained to PI-RADS ≤ 2 and 'difficult' cases. Six examinations in months 1–6 were not assigned a PI-RADS v2 score because the imaging was performed following definitive local therapy. Upper and lower CI refer to 95% confidence intervals.

reviewed during months 7–12 were restricted to 'difficult' and 'negative' cases because the primary weakness detected collaboratively during months 1–6 was false-negative examinations (i.e., examinations scored PI-RADS ≤ 2 but with Gleason ≥ 7 cancer on histology) at the in-training site (Fig. 4).

The positive predictive values for PI-RADS scores of 3–5 were similar at the in-training and supervising

Fig. 4. False-negative examinations (+) reported by the in-training institution plotted over time (Quarter 1: $n = 7$; Quarter 2: $n = 7$; Quarter 3: $n = 1$; Quarter 4: $n = 3$). False-negative is defined as PI-RADS 1–2 assigned by the in-training institution and Gleason ≥ 7 prostate cancer identified on histology.

Table 3. Positive predictive values for Gleason ≥ 7 prostate cancer stratified by maximum PI-RADS score

Maximum PI-RADS score and histology	In-training institution (%)	Supervising institution (%)
PI-RADS 3	25 (3/12)	29 (5/17)
PI-RADS 4	64 (14/22)	62 (16/26)
PI-RADS 5	75 (9/12)	72 (18/25)

Histology is the most aggressive Gleason score obtained by any method (i.e., standard biopsy, targeted biopsy, radical prostatectomy). Positive predictive values only include patients with available histology

institutions (Table 3). Fourteen percent (11/80) of PI-RADS ≥ 3 examinations were reclassified PI-RADS ≤ 2 (Table 4); one had post-MR histology and none were diagnosed with Gleason ≥ 7 prostate cancer. Nineteen percent (30/162) of PI-RADS ≤ 2 examinations were reclassified PI-RADS ≥ 3; of these, 23 had post-MRI histology and 57% (13/23) had clinically significant cancer (Table 4).

Thirteen [5.2% (13/249)] Gleason ≥ 7 prostate cancers were identified by the supervising institution that were not identified by the in-training institution (Table 4). False-negative examinations ($N = 18$) were more common at the in-training institution during the first six months compared to the second six months [9% (14/160) vs. 4% (4/89)] (Fig. 4). Of the 4 false-negative results that occurred in the second six months, 3 (75%) were by a radiologist who did not participate in months 1–6 {i.e., only 1 false-negative result [1% (1/67)] during months 7–12 was attributable to one the three radiologists who participated in months 1–6}. Five false-negative results occurred at the supervising institution; all were Gleason $3 + 4 = 7$ (Table 4).

Inter-rater agreement by quarter for the in-training and supervising interpretations was excellent for calculated prostate volume, substantial to almost perfect for subjective volume of BPH, moderate to substantial for degree of median lobe enlargement, poor to fair for membranous urethra length, and fair to substantial for T3 disease (Table 2).

Table 4. Histologic outcomes by site and maximum PI-RADS score

Maximum PI-RADS score and histology	In-training institution	Supervising institution	Difference
PI-RADS ≤ 2			
Total	162	143	− 19
No histology	118	121	
Histology available	44	22	
Benign	11	9	− 2
Gleason 3 + 3	15	8	− 7
Gleason 3 + 4	15	5	− 10
Gleason 4 + 3 *or higher*	3	0	− 3
PI-RADS 3			
Total	15	22	7
No histology	3	17	
Histology available	12	5	
Benign	4	5	1
Gleason 3 + 3	5	7	2
Gleason 3 + 4	3	5	2
Gleason 4 + 3 *or higher*	0	0	0
PI-RADS 4			
Total	38	36	− 2
No histology	16	10	
Histology available	22	26	
Benign	2	2	0
Gleason 3 + 3	6	8	2
Gleason 3 + 4	8	10	2
Gleason 4 + 3 *or higher*	6	6	0
PI-RADS 5			
Total	27	41	14
No histology	15	16	
Histology available	12	35	
Benign	2	3	1
Gleason 3 + 3	1	4	3
Gleason 3 + 4	3	9	6
Gleason 4 + 3 *or higher*	6	9	3

Histology is the most aggressive Gleason score obtained by any method (i.e., standard biopsy, targeted biopsy, radical prostatectomy)

Discussion

The 12-month longitudinal long-distance quality assurance program presented here is an example of how academic and private facilities without an official relationship can partner to promote high-quality imaging, interpretation, and patient care during the adoption of a new complex imaging technology. Based on our experience, a mentoring period of approximately 6 months encompassing at least 150 MRIs (~ 50 MRIs/radiologist) is recommended for radiology practices interested in developing a new prostate

MRI program. The primary weaknesses of the in-training site during the first 1–6 months were image quality impairment and false-negative examinations. Through iterative improvement of image quality and mentored case review, we saw substantial improvements in image quality, inter-rater agreement, and diagnostic accuracy. In months 7–12, when only PI-RADS ≤ 2 and 'difficult' cases were reviewed, the false-negative results for Gleason ≥ 7 prostate cancer were 4% (4/89) overall, and only 1% (1/67) for the radiologists who participated for the entire relationship. The decline in maximum PI-RADS v2 score agreement observed in quarter 4 [ICC: 0.39 (0.11–0.61)] is likely multifactorial; it is probably related to a combination of adding a fourth in-training radiologist and restricting case review to only 'difficult' and PI-RADS ≤ 2 examinations.

Other studies have investigated the learning curve associated with prostate MRI [5–10]. Gaziev et al. [5] in 2016 conducted a single-center longitudinal analysis of 340 men who underwent prostate MRI followed by transperineal fusion biopsy and found that the diagnostic performance of MRI and biopsy naturally improved over a 22-month period. There was no formal educational process described. Rosenkrantz et al. [6] in 2017 conducted a single-center retrospective study of 6 second-year radiology residents interpreting 124 prostate MRI examinations, half with and half without immediate feedback, and found that all learners experienced a rapid improvement in diagnostic accuracy that began to plateau after 40 examinations. Akin et al [8] in 2010 conducted a single-center prospective study of 11 radiology fellows exposed to weekly interactive tutorials and a set of didactic lectures and found that performance improved over a 20-week period, particularly with regard to tumor localization and extracapsular extension. Garcia-Reyes et al. [10] in 2015 conducted a single-center retrospective study of 5 radiology fellows analyzing 31 prostate MRI examinations and found that after being exposed to two lectures, there were at least short-term improvements in diagnostic accuracy. Each of these studies were conducted in an academic environment with direct in-person interventions. Such interventions likely do not directly apply to radiology practices starting a new prostate MRI program, where imaging protocol development and rapid acquisition of new skills in the absence of mentored oversight are key challenges. In such settings, long-distance mentored quality assurance may be a better option.

In addition to demonstrated gains in image quality and diagnostic performance, other interesting themes emerged over the duration of our intervention. Inter-rater agreement for triplanar calculated prostate volume was excellent (ICC: 0.92–0.95) and inter-rater agreement for subjective measures of BPH volume was substantial to almost perfect. Both of these manual methods appear to be consistent ways of conveying the size of the prostate.

Membranous urethra length has been shown in a meta-analysis [16] to predict the likelihood of incontinence after prostatectomy; however, in our study, we found that inter-rater agreement for this assessment between the in-training and supervising sites was poor to fair, indicating that if this measurement is to be useful clinically, more reliable methods of measuring are needed. Finally, inter-rater agreement for T3 disease (i.e., extracapsular extension or seminal vesicle invasion) was wide-ranging—from fair to substantial. This may be because the sample was not restricted to subjects with high-risk disease, preventing repeated exposure to this determination.

Our report has limitations. This was a prospective quality assurance effort and not all MRI examinations have histologic correlation. This is particularly true for those assigned a PI-RADS score of 2, which in both practices was commonly interpreted by urologists to indicate a low likelihood of clinically significant cancer. Therefore, negative predictive values, sensitivity, and specificity cannot be calculated reliably; however, we did observe a substantial decrease in false-negative results during months 7–12. Because we restricted the types of cases reviewed during months 7–12, it is difficult to determine the long-term durability of our results. The positive predictive values we observed by PI-RADS score are likely higher than what would be observed in a normal context because the patient population included some patients being imaged for pre-operative planning ($N = 36$) and some patients being imaged following local therapy for prostate cancer ($N = 7$). These data reflect the evaluations by 4 in-training radiologists and 1 supervising radiologist. It is possible that the results might be different if other radiologists were involved. We minimized the biases intrinsic to assessments of inter-rater agreement by ensuring that all reporting at the in-training and supervising sites was done using the same prostate MRI reporting template. Finally, the in-training and supervising sites were both motivated to participate in the program. If such long-distance quality assurance was conducted in a heavy-handed or mandatory fashion, it may be less well received.

In conclusion, remote quality assurance of prostate MRI is feasible and useful, and can enable new programs to gain durable skills with minimal risk to patients. Academic institutions with an educational mission and experience in prostate MRI can partner with other radiology practices to train inexperienced radiologists, optimize prostate MRI acquisition, and deliver high-quality interpretations. The expansion and reliability of prostate MRI is directly dependent on the methods of training in-practice radiologists. The training experience we present resulted in marked improvement in image acquisition and interpretation over a 12-month period and can be considered a model for post-graduate training in new complex radiology techniques.

References

1. Weinreb JC, Barentsz JO, Choyke PL, et al. (2016) PI-RADS prostate imaging-reporting and data system: 2015, version 2. Eur Urol 69:16–40
2. Vargas HA, Hotker AM, Goldman DA, et al. (2016) Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference. Eur Radiol 26:1606–1612
3. Siddiqui MM, Rais-Bahrami S, Turkbey B, et al. (2015) Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. JAMA 13:390–397
4. Rosenkrantz AB, Ginocchio LA, Cornfeld D, et al. (2016) Interobserver reproducibility of the PI-RADS version 2 lexicon: a multicenter study of six experienced prostate radiologists. Radiology 280:793–804
5. Gaziev G, Wadhwa K, Barrett T, et al. (2016) Defining the learning curve for multiparametric magnetic resonance imaging (MRI) of the prostate using MRI-transrectal ultrasonography (TRUS) fusion-guided transperineal prostate biopsies as a validation tool. BJU Int 117:80–86
6. Rosenkrantz AB, Ayoola A, Hoffman D, et al. (2017) The learning curve in prostate MRI interpretation: self-directed learning versus continual reader feedback. AJR Am J Roentgenol 208:W92–W100
7. Harris RD, Schned AR, Heaney JA (1995) Staging of prostate cancer with endorectal MR imaging: iessons from a learning curve. RadioGraphics 15:813–829
8. Akin O, Riedl CC, Ishill NM, et al. (2010) Interactive dedicated training curriculum improves accuracy in the interpretation of MR imaging of prostate cancer. Eur Radiol 20:995–1002
9. Latchamsetty KC, Borden LS Jr, Porter CR, et al. (2007) Experience improves staging accuracy of endorectal magnetic resonance imaging in prostate cancer: what is the learning curve? Can J Urol 14:3429–3434
10. Garcia-Reyes K, Passoni NM, Palmeri ML, et al. (2015) Detection of prostate cancer with multiparametric MRI (mpMRI): effect of dedicated reader education on accuracy and confidence of index and anterior cancer diagnosis. Abdom Imaging 40:134–142
11. Mager R, Brandt MP, Borgmann H, et al. (2017) From novice to expert: analyzing the learning curve for MRI-transrectal ultrasonography fusion-guided transrectal prostate biopsy. Int Urol Nephrol. https://doi.org/10.1007/s11255-017-1642-7
12. Kierans AS, Taneja SS, Rosenkrantz AB (2015) Implementation of multi-parametric prostate MRI in clinical practice. Curr Urol Rep 16:56
13. Leake JL, Hardman R, Ojili V, et al. (2014) Prostate MRI: access to and current practice of prostate MRI in the United States. J Am Coll Radiol 11:156–160
14. Muthigi A, Sidana A, George AK, et al. (2017) Current beliefs and practice patterns among urologists regarding prostate magnetic resonance imaging and magnetic resonance-targeted biopsy. Urol Oncol 35:32.e1–32.37.
15. Alpert HR, Hillman BJ (2004) Quality and variability in diagnostic radiology. J Am Coll Radiol 1:127–132
16. Mungovan SF, Sandhu JS, Akin O, et al. (2007) Preoperative membranous urethra length measurement and continence recovery following radical prostatectomy: a systematic review and meta-analysis. Eur Urol 71(3):368–378
17. Rosenkrantz AB, Shanbhoque AK, Wang A, et al. (2016) Length of capsular contact for diagnosing extraprostatic extension on prostate MRI: assessment at an optimal threshold. J Magn Reson Imaging 43:990–997
18. Viera AJ, Garrett JM (2005) Understanding interobserver agreement: the kappa statistic. Fam Med 37:360–363
19. Hallgren KA (2012) Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 8:23–34