bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Deep-Learning-Enabled Differentiation between Intraprostatic Gold Fiducial Markers and Calcification in Quantitative Susceptibility Mapping

Ashley Wilton Stewart[1], Jonathan Goodwin[2,3], Matthew Richardson[2], Simon Daniel Robinson[4,5], Kieran O'Brien[5,6,7], Jin Jin[5,6,7], Markus Barth[1,5,7], Steffen Bollmann[1,5,7,8]

[1]School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Australia
[2]Department of Radiation Oncology, Calvary Mater Hospital, Newcastle, New South Wales, Australia
[3]School of Information and Physical Sciences, University of Newcastle, Newcastle, New South Wales, Australia
[4]High Field MR Center, Department of Biomedical Imaging and Image-Guided Therapy, Medical University of Vienna, Vienna, Austria
[5]Centre for Advanced Imaging, The University of Queensland, Brisbane, Australia
[6]Siemens Healthcare Pty Ltd, Brisbane, Australia
[7]ARC Training Centre for Innovation in Biomedical Imaging Technology, The University of Queensland, Brisbane, Australia
[8]Queensland Digital Health Centre, The University of Queensland, Brisbane, Australia

Corresponding author: Ashley Stewart; ashley.stewart@uq.edu.au. @stewart_au

Word count: 3,347 words

Submitted to Magnetic Resonance in Medicine.

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Purpose: Interest is growing in MR-only radiotherapy (RT) planning for prostate cancer (PCa) due to the potential reductions in cost and patient exposure to radiation, and a more streamlined workflow and patient imaging pathway. However, in MRI, the gold fiducial markers (FMs) used for target localization appear as signal voids, complicating differentiation from other void sources such as calcifications and bleeds. This work investigates using Quantitative Susceptibility Mapping (QSM), an MRI phase post-processing technique, to aid in the differentiation task. It also presents deep learning models that capture nuanced information and automate the segmentation task, facilitating a streamlined approach to MR-only RT.

Methods: CT and MRI, including GRE and T1-weighted imaging, were acquired from 26 PCa patients, each with three implanted gold FMs. GRE data were post-processed into QSM, $T2^*$, and $R2^*$ maps using QSMxT's body imaging pipeline. Statistical analyses were conducted to investigate the quantitative differentiation of FMs and calcification in each contrast. 3D U-Nets were developed using fastMONAI to automate the segmentation task using various combinations of MR-derived contrasts, with a model trained on CT used as a baseline. Models were evaluated using precision and recall calculated using a leave-one-out cross-validation scheme.

Results: Significant differences were observed between FM and calcification regions in CT, QSM and $T2^*$, though overlap was observed in QSM and $T2^*$. The baseline CT U-Net achieved an FM-level precision of $\approx$ 98% and perfect recall. The best-performing QSM-based model achieved precision and recall of 80% and 90%, respectively, while conventional MRI had values below 70% and 80%, respectively. The QSM-based model produced segmentations with good agreement with the ground truth, including a challenging FM that coincided with a bleed.

Conclusion: The model performance highlights the value of using QSM over indirect measures in MRI, such as signal voids in magnitude-based contrasts. The results also indicate that a U-Net can capture more information about the presentation of FMs and other sources than would be possible using susceptibility quantification alone, which may be less reliable due to the diverse presentation of sources across a patient population. In our study, QSM was a reliable discriminator of FMs and other sources in the prostate, facilitating an accurate and streamlined approach to MR-only RT.

Keywords: Quantitative susceptibility mapping, QSM, MR-only radiotherapy, MR-guided radiotherapy, prostate cancer, fiducial markers, gold fiducial markers, deep learning, medical imaging, medical image segmentation, U-Net.

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 Introduction

Prostate cancer (PCa) is the second most common cancer in men, with over 1.2 million new cases and 350,000 deaths annually [1]. For localized PCa, treatments include surgery and various forms of radiotherapy (RT), with intensity-modulated RT (IMRT) emerging as the standard of care, primarily due to its ability to target tumor volumes while minimizing doses to surrounding tissues precisely. However, geometric misses can occur due to inter- and intra-fractional patient motion [2, 3]. Such misses can be mitigated using gold fiducial markers (FMs) implanted into the prostate gland under ultrasound guidance for image-matching purposes [4, 5]. In the conventional CT-MRI planning workflow, MRI is used for its superior soft-tissue contrast in target delineation. Meanwhile, CT provides the attenuation information necessary for dosimetry and FM visualization. In the treatment room, the available imaging modality, typically kilovoltage or megavoltage X-ray, is used to verify the target positioning by matching the positions of FMs in the co-registered planning images. Recently, interest in MR-only RT planning has grown [6–8]. An MR-only workflow uses planning images derived solely from MRI rather than the conventional CT-MRI workflow. This approach offers several benefits, including reduced geometric uncertainty by eliminating the need for CT-MR registration, reduced cost, a more streamlined patient imaging pathway, and reduced patient exposure to ionizing radiation.

Despite the potential advantages of MR-only RT planning, the approach poses several challenges. Conventional MRI's non-quantitative nature complicates dose calculations, an issue now addressed by synthetic CT (sCT) derived from MRI, which aligns closely with CT-based planning [9–12]. Geometric distortions in MRI, especially at air-tissue boundaries, affect spatial accuracy and remain a challenge [13, 14]. A major concern remains in the detection of FMs in MRI. In MRI, FMs appear as signal voids, complicating differentiation from other void sources such as calcifications and bleeds.

Since differentiating void sources in the prostate can be challenging and time-consuming, a range of automated FM detection techniques have recently been proposed for MR-only RT, with initial approaches based on template matching. An early method of this type used spectral clustering of FM templates constructed from manual contours, achieving FM detection accuracies of 63% and 88% for $T2^*$- and T1-weighted GRE sequences, respectively, over 15 patients [15]. Expanding on this technique, another study applied simulated complex-valued templates, reporting a true positive rate (TPR) of 96% across 17 patients [16]. Both investigations highlighted the challenges faced in distinguishing FMs from other sources such as calcifications and bleeds. Another method used logistic regression to integrate features across multi-parametric MRI and achieved a TPR of 94% over 32 patients [17]. A similar approach cross-referenced data from various MRI sequences to pinpoint FM candidates in T2w images, resulting in a TPR of 84% across 40 patients, with most false positives due to calcifications [18]. A significant breakthrough came with the adoption of deep learning techniques, where one study showcased a detection sensitivity of 97.4% across 39 prostate cancer patients, matching human expert performance in the same images [19].

Quantitative susceptibility mapping (QSM) is an MRI post-processing technique which calculates the magnetic susceptibility distribution of imaged objects from the static magnetic field [20–23]. Since QSM is sensitive to susceptibility variations, it has the potential to identify and differentiate FMs

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

from calcification quantitatively. To date, QSM has mainly been applied in human brain imaging applications [24]. Investigations of QSM in the body present several challenges. First, the chemical shift effect of protons in fat can cause phase discontinuities and geometric distortions at fat-water interfaces [25], which gives rise to unwrapping and other artifacts in QSM. Second, the higher dynamic range of susceptibility values in body regions is problematic for regularizers optimized for narrower ranges in brain imaging. However, QSM investigations in the body are rapidly expanding with the development of advanced water-fat separation techniques, robust QSM reconstruction pipelines, and artifact reductions [26–31].

Initial work in prostate QSM investigated the visualization of post-biopsy calcification to aid in PCa diagnosis [32]. Realizing the potential for QSM to enable MR-only RT planning, later studies investigated QSM for localizing calcification as a natural alternative to implanted FMs [33], a concept initially applied in cone-beam CT [34]. Concurrent work specifically investigated using QSM to differentiate FMs from calcification *in silico* [35], including paramagnetic low dose-rate (LDR) brachytherapy seeds for post-implant dosimetry [36]. The same authors later expanded their work in LDR seeds to three patients for *in vivo* validation [37]. These LDR markers are paramagnetic and appear hyperintense in QSM, making them straightforward to differentiate from diamagnetic calcifications appearing hypointense. However, the gold FMs conventionally used for prostate IGRT are diamagnetic, and differentiation in QSM will rely more heavily on quantification or complementary MRI contrasts. Recent work has demonstrated that QSM may provide sufficient information for the differentiation task with a dataset of 7 PCa patients [38]. However, FM identification or differentiation requires validation using larger datasets and is yet to be automated.

This study investigates the effectiveness of a QSM post-processing pipeline optimized for body imaging applications in identifying and differentiating FMs from calcifications in a dataset of 26 PCa patients. Initially, the differentiation capability of multiple MRI contrasts, including QSM and $T2^*$ maps, is quantitatively assessed through statistical testing. Subsequently, 3D U-Nets are developed to automate the classification and segmentation of each region. The models are trained using combinations of MRI contrasts, including QSM, $R2^*$, SWI, T1-weighted, and GRE magnitude images, to identify which combinations provide the most valuable information for the segmentation task. By mitigating the imaging challenges posed by calcifications in the presence of FMs, this work facilitates a more accurate and streamlined approach to MR-only RT.

2 Methods

2.1 Data and Acquisition

MRI and CT were acquired from 26 PCa patients, each with three implanted Gold Markers (Riverpoint Medical, Portland, Oregon, United States), cylindrical and 1×3 mm. All data were acquired with the approval of the relevant ethics committees. Given the study's focus on differentiating FMs and calcifications, PCa patients with calcifications were preferentially scanned. In this sample, calcifications were present in 14 patients ($\approx 54\%$).

CT data were acquired using a Siemens SOMATOM Confidence CT scanner (Siemens Healthcare, Erlangen, Germany) on software version VB10A (imaging parameters: imaging volume superior

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

scan limit included L5 to lesser trochanters inferiorly, imaging resolution 0.98 × 0.98 × 2 mm, an image matrix size of 512 × 512 × 188, and kVp=120 kV).

MRI data were acquired using a MAGNETOM Skyra 3T scanner (Siemens Healthcare, Erlangen, Germany) with a 32-channel spine coil and 18-channel body coil on software version VE11. Acquisitions include T1-weighted anatomical scans and a gradient-echo (GRE) sequence used for QSM reconstruction. The GRE used a bipolar readout scheme and 1.4 mm^3 isotropic resolution, TEs=2.46/4.92/7.39/9.84/12.3/14.8/17.22 ms, TR=25 ms, TA=2:54 min:sec, and FA=45°. Phased array channels were combined using the pre-scan normalize plus adaptive combine approach [39]. Due to the chemical-shift effect causing opposing geometric shifts between odd and even echoes, only the even, in-phase echoes were kept for further post-processing.

2.2 Data post-processing

GRE images were processed using QSMxT v5.1.0 [30, 40] using the default settings for body imaging applications. The body imaging pipeline applies TGV-QSM [41, 42] for the phase unwrapping, background field correction, and dipole inversion steps, an automated threshold-based masking algorithm, and a two-pass artifact-reduction algorithm. QSMxT also produced susceptibility-weighted imaging (SWI) and $T2^*/R2^*$ maps using CLEAR-SWI [43] and NumART2star [44], respectively, from the MriResearchTools library [45].

CT images were registered and resampled to the GRE space using ITK-SNAP [46]. This resampling enabled a uniform U-Net training procedure by eliminating the need for architecture modifications across the modalities, ensuring a fair comparison of learning potential at the MRI resolution. The authors roughly segmented FMs, calcification, and prostate tissue regions separately for the registered CT and GRE images. To ensure a consistent segmentation approach, the rough segmentations were automatically filtered. The GRE FM segmentations were filtered to include susceptibilities three standard deviations below the mean prostate susceptibility, while calcifications were constrained by two standard deviations. In the CT images, FM segmentations were filtered to include values ten standard deviations above the prostate mean attenuation, while calcifications were constrained by four standard deviations. These thresholds were determined empirically, ensuring segmentations covered the distinct FM and calcification regions without including surrounding tissues. Segmentations were further refined using Gaussian smoothing and morphological operations.

The presentation of calcifications can vary considerably compared with FMs. Calcifications can be large or small, with varied levels of concentration. While the U-Net models were trained using all calcification segmentations, statistical analyses included only calcifications that were of similar size to FMs and which could be considered challenging FM candidates. This final filtering step included only calcifications within 3 standard deviations of the mean FM size in voxels.

2.3 Model and training

A 3D U-Net was implemented using fastMONAI [47, 48], which interfaces with fastai [49], MONAI [50], and PyTorch [51]. The U-Net was based on the fastMONAI example for multi-class segmentation, which comprises five levels with channels (16, 32, 64, 128, 256) and strides (2, 2, 2, 2), along

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with two residual units in each level. The U-Net was adapted to take multi-modal images as inputs and output predictions for background, calcification, and FM regions. Models were trained on various combinations of images to investigate how they impact performance. These combinations include standalone modalities of CT, QSM, SWI, T1, $R2^*$, and GRE (combined across echoes using the sum-of-squares), as well as combinations of (QSM, T1, and $R2^*$), (QSM and T1), and (QSM and SWI).

For training, input images and masks were initially cropped to $80 \times 80 \times 80$ voxels, and data were processed in batches of six. The training was configured for 800 epochs, with early stopping triggered after 200 epochs of no improvement. The datasets were augmented using random flipping along the left-right and anterior-posterior axes, random affine transformations with rotations up to 90 degrees, and Z-normalization. To handle possible intraprostatic bleeds causing FMs to appear hyperintense in QSM, random multiplications of -1 were made to the FM QSM values in 5% of training examples. The model was optimized using the RAdam optimizer with the Lookahead method, known as Ranger, and an initial learning rate of 0.003 was set. A cosine annealing learning rate schedule was employed during the training process. The loss function combined Dice loss and cross-entropy loss with one-hot encoding applied to the target masks.

For evaluation, a leave-one-out cross-validation (LOOCV) scheme was implemented. This approach required training 26 models or *folds*, each with a different validation example. This approach allows for a more reliable assessment of model performance without sacrificing valuable training examples for a separate testing set. No hyperparameter tuning was conducted to minimize the risk of overfitting to the validation examples. Each model's performance was evaluated by averaging metrics across the 26 folds. Metrics for evaluation are the multi-class Dice score, voxel-level true positive and false positive rates, voxel-level recall and precision, and FM-level recall and precision. The multi-class Dice score quantifies the overlap between predicted and ground truth segmentations. Precision and recall reflect the model's ability to correctly identify FMs, differentiate them from other sources, and recall actual FMs reliably. Voxel-level metrics are calculated for individual voxels, whereas FM-level metrics are calculated across whole FMs.

3 Results

78 FMs were observed across each contrast and modality, corresponding to the expected 3 FMs per patient. FMs appeared hyperintense in CT and hypointense in T1-weighted, SWI, and GRE magnitude images. In QSM, 75 FMs appeared hypointense, one appeared hyperintense, and the remaining two appeared isointense (see Figure 1). A radiologist confirmed the hyperintensity as a bleed at the FM insertion site. The two isointensities occurred in one patient and coincided with hyperplasic tissue due to benign prostatic hyperplasia (BPH) and a possible bleed. The $R2^*$ images were noisy, and FMs were difficult to discern but tended towards hyperintense compared to the surrounding prostate tissue.

Statistical significance was observed across the distributions of mean FM and calcification values in the quantitative images, including CT, QSM, and $T2^*$ (see Figure 2). There was no overlap between the calcification and FM mean value distributions in CT histograms, though overlap was observed

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

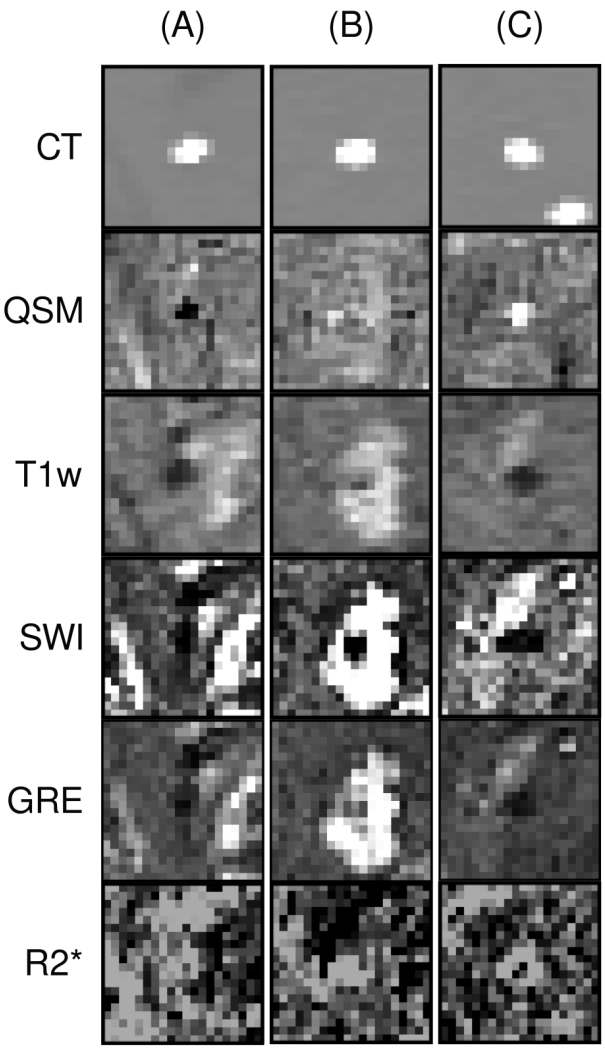


Figure 1: Presentation of FMs across image types; cropped axial regions. A) depicts a typical FM appearing hypointense in QSM; B) depicts an FM appearing within hyperplasic tissue and a small bleed which appears isointense in QSM; C) depicts an FM coinciding with a bleed which appears hyperintense in QSM. The typical FM appearance in QSM has a clearer boundary than in the other MR-derived images.

in QSM, with more extensive overlap in $T2^*$.

The performance of each U-Net was evaluated using Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves. Curves were computed based on the FM labeling of individual voxels, with the metrics averaged across cross-validation folds. For the ROC curves (see Figure 3), CT had the highest performance with an area-under-curve (AUC) of 0.99, followed closely by the QSM model (AUC of 0.97). More conventional MRI trailed behind the models incorporating QSM, with AUCs ≤ 0.87. CT again demonstrated superior performance in the PR curves with AUC $= 0.91$ (see Figure 4). Models incorporating QSM were clustered with AUC ≈ 0.65, with other MRI formed a separate cluster with inferior AUCs ≈ 0.3.

Precision and recall metrics were also computed at the FM level across models and averaged across cross-validation folds (see Table 1). The CT model achieved a precision of 0.98 ± 0.07 and a perfect

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

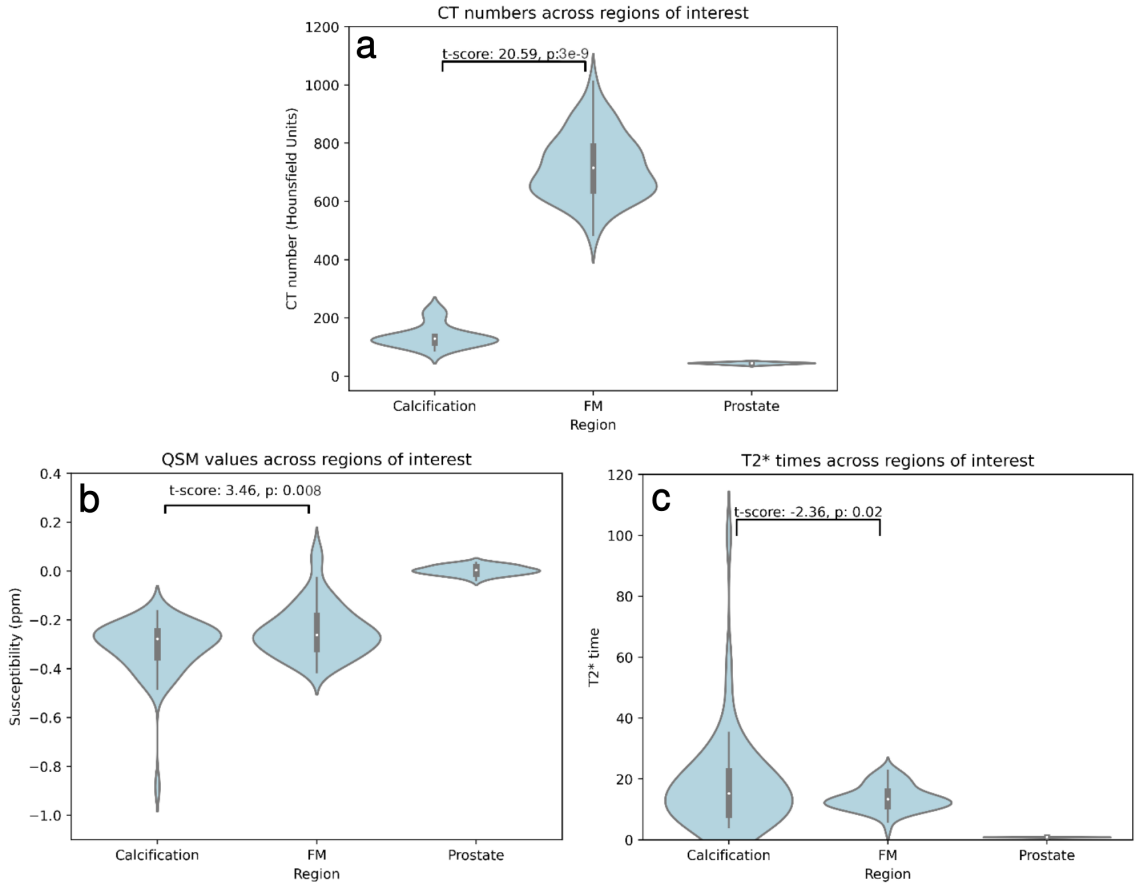


Figure 2: Region-level mean values for calcifications, FMs, and prostate tissue in A) CT numbers; B) QSM values; and C) $T2^*$ times. Calcifications were filtered to those with sizes within three standard deviations of the mean FM size. Violin widths are equal. The t-scores and associated p-values for the calcification and FM regions are annotated. While there is significant differentiation between the means of the FM and calcification distributions in all contrasts, there is some overlap in QSM and $T2^*$.

recall. The QSM model achieved a precision of 0.8 ± 0.21 and recall 0.9 ± 0.2, similar to the other models incorporating QSM. Models trained using other MR-derived images, including SWI, T1, $R2^*$ maps, and GRE magnitude images, had achieved precision ≤ 0.67 and recall ≤ 0.79, along with wider standard deviations.

Representative QSM images depicting FMs and calcification were visualized along with ground truth segmentations, model predictions using the QSM U-Net, and resultant labels (see Figure 5). The model's predictions are closely aligned with the ground truth labels.

4 Discussion

This study evaluated the efficacy of a QSM post-processing pipeline for distinguishing intraprostatic FMs from calcifications in 26 PCa patients in conjunction with various MRI contrasts and automated the segmentation task with a 3D U-Net, facilitating a streamlined MR-only RT approach.

Our initial statistical analysis showed significant differences in mean CT numbers, QSM values, and

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

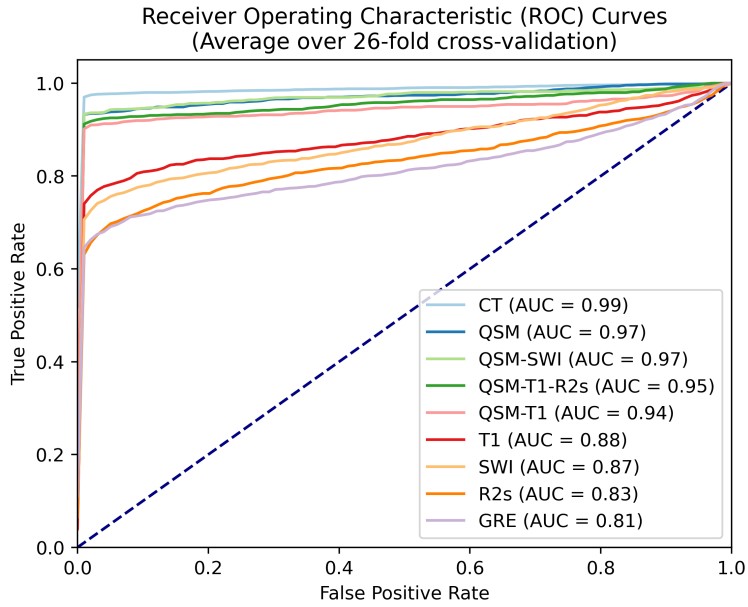


Figure 3: Voxel-level Receiver Operating Characteristic (ROC) curves for each imaging modality on validation data and evaluated against the FM label, averaged over all cross-validation folds.

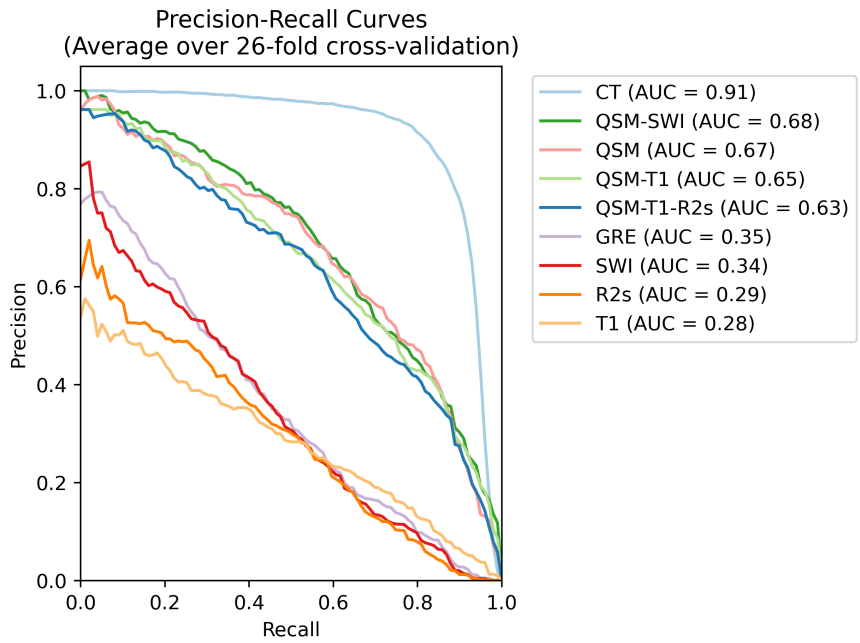


Figure 4: Voxel-level Precision-Recall (PR) curves for each imaging modality on validation data and evaluated against the FM label, averaged over all cross-validation folds.

$T2^*$ times of distinct FM and calcifications. While there was no overlap between these regions in CT numbers, the QSM values and $T2^*$ times had some overlap. This indicates that discriminating the sources solely on MR-based quantitative values remains challenging despite the significant differences in their distributions.

U-Nets were trained for FM and calcification segmentation using single or multi-contrast input. The

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Model	Precision	Recall
CT	0.98 ± 0.07	1 ± 0
QSM-T1-R2s	0.81 ± 0.25	0.9 ± 0.25
QSM-T1	0.77 ± 0.25	0.87 ± 0.28
QSM-SWI	0.8 ± 0.2	0.93 ± 0.19
QSM	0.8 ± 0.21	0.9 ± 0.2
T1	0.67 ± 0.31	0.79 ± 0.34
SWI	0.57 ± 0.28	0.65 ± 0.35
R2s	0.61 ± 0.27	0.71 ± 0.31
GRE	0.56 ± 0.31	0.59 ± 0.36

Table 1: FM-level precision and recall across models, averaged over cross-validation folds.

standalone CT model excelled in identifying FMs with 98% precision and perfect recall. Meanwhile, QSM-based models showed commendable performance, with the standalone QSM model achieving 80% precision and 90% recall. In contrast, models trained on MRI without QSM performed poorly, with precision and recall values below 70% and 80%, respectively.

Prior work investigating FM localization using QSM demonstrated quantitative differentiation of FMs from calcifications [38]. The authors measured FMs at -31.5 ± 2.0 ppm and calcifications at -14.6 ± 0.9 ppm in a dataset of seven PCa patients, two of which were reported to have calcifications. Our study did not replicate these findings, with extensive overlap in mean susceptibility values between FMs and calcifications across 26 PCa patients. However, a direct comparison is challenging because of the small number of calcifications in the cited work and a lack of clarity on FM and calcification segmentation methods. Moreover, calcification presentation can vary considerably in size and concentration, and the material and size of the FMs used could have a substantial influence on findings. Our analysis considered the mean values of distinct calcified regions, similar in size to the 1×3 mm pure gold FMs, segmented semi-automatically as described in the methods section.

Most automated FM detection techniques in existing literature rely on magnitude images [15, 18, 19], with some including multi-parametric MRI [17, 18], and another using complex-domain images [16]. The visualization of FMs as distinct signal voids arises from their underlying magnetic susceptibility difference from prostate tissues. This insight motivated our investigation into techniques based on QSM images, which offer a more direct susceptibility measurement. Although the best-performing technique to date uses magnitude images alone [19], it is worth noting that the cited study used larger FMs than the present study, which could explain their detection efficacy. In the present study, several U-Nets were trained using the multi-parametric approach with combinations of MR images including (QSM, T1, and $R2^*$), (QSM and T1), and (QSM and SWI) to determine whether they could enhance a deep-learning-based solution. As hypothesized, the U-Nets trained on QSM images surpassed the performance of those trained on T1, R2, SWI, and GRE magnitude images. However, combining multiple contrasts did not markedly improve performance in our study. In some instances, performance even deteriorated slightly, possibly due to the introduction of noise or redundant information that detracted from the distinguishing features offered by QSM alone.

The performance of the U-Nets may be improved further through hyperparameter tuning. Avoiding

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

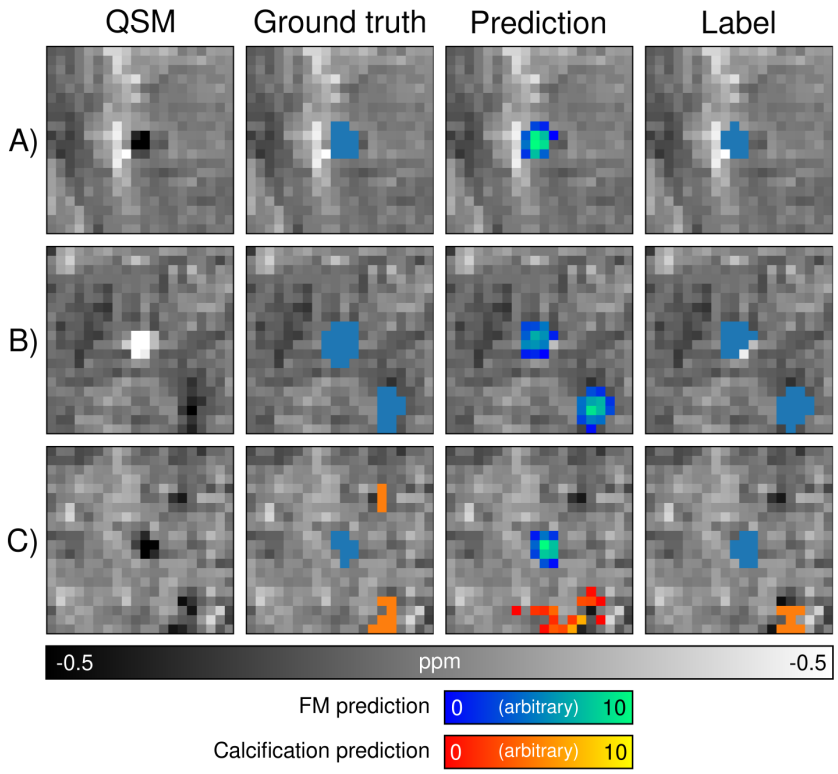


Figure 5: Representative examples of FMs and calcification in QSM, including ground truth segmentation labels, predictions produced by the segmentation model trained on QSM only (in arbitrary units), and the predicted label based on the maximum prediction across each class. A) depicts a typical FM; B) depicts an FM appearing hyperintense due to a bleed at the insertion site; C) depicts an FM with nearby calcification.

this in our study was a deliberate choice to prevent overfitting to the limited dataset. The LOOCV approach facilitates a fair evaluation of the U-Nets and maximizes training data availability while reserving only one example for evaluation in each of the 26 folds. However, a more extensive dataset may facilitate a separate testing set and permit hyperparameter tuning to optimize performance on validation data before final testing. Another reason hyperparameter tuning was avoided was to ensure the same network architecture could be applied for each U-Net regardless of the input image combinations. However, each image type or combination might require individualized hyperparameter adjustments to perform optimally.

Another way the performance of the U-Nets may be improved is by tweaking the acquisition. The current best-performing FM detection technique used an acquisition with a slice thickness of 2.8 mm and an in-plane resolution of 1.46 mm [19], gathering more signal per voxel as in our isotropic resolution of 1.4 mm^3. The isotropic voxel size was chosen because it improves quantification in QSM as recommended in the recent QSM consensus paper [52]. Investigating sequence optimizations, such as altering the slice thickness, may improve model performance and visualization of sources in QSM for this particular application.

Multiple studies have noted that differentiation of FMs from other sources, such as bleeds and calcification, can be challenging in magnitude images [15, 16, 18]. This generally occurs with calci-

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

fications that result in a signal void similar in size and shape to those caused by FMs. Previously proposed automated methods have resulted in false negatives where bleeds occur at the FM insertion site. Although our dataset contains only a single instance of such a bleed, the data augmentations that simulate bleeds aided the QSM U-Net in identifying the FM successfully. This outcome highlights the benefits of realistic simulated datasets and augmentations, which are becoming increasingly important for clinical deep-learning tools where large-scale datasets are challenging to acquire [53, 54].

One major limitation of the present study is the dataset's relatively small size. While our dataset of 26 PCa patients is more extensive than similar studies, it may not capture the anatomical variability of a broader population. Uncommon cases were scarce, such as confirmed bleeds at FM insertion sites, instances of BPH, or other gross pathology causing complex tissue composition changes. This paucity could affect the U-Net's ability to identify FMs at clinically essential edge cases effectively. Therefore, in addition to realistic simulations and data augmentations, acquiring larger datasets over broad populations will be crucial to understanding these edge cases and developing more robust models.

Future research should aim to build larger datasets encompassing a more diverse cohort of PCa patients. This expansion would improve model performance, generalizability, and robustness across patient subsets. The current findings in uncommon cases like bleeds emphasize the importance of obtaining more such data for richer insights into their QSM and conventional MRI interpretations. While our simulated bleeds approach proved effective, there is potential to leverage advanced augmentation techniques further. As findings evolve, assessing the practical utility of automated FM detection tools in clinical contexts will become crucial, comparing their outcomes with manual radiologist annotations to determine clinical readiness.

5 Conclusions

This study evaluates the potential of QSM to identify and distinguish intraprostatic FMs from calcifications in 26 PCa patients. The developed U-Nets subsequently automate the segmentation task. The model performance highlights the value of using QSM over indirect measures such as signal voids in magnitude-based contrasts and SWI. The results also indicate that a U-Net can capture more information from the presentation of FMs and other sources than would be possible using susceptibility quantification alone, which may be less reliable due to the diverse presentation of sources across a patient population. In our study, QSM was a reliable discriminator of FMs and other sources in the prostate, facilitating an accurate and streamlined approach to MR-only RT. Future work should focus on expanding training datasets to capture the variability across wider patient populations and develop clinical tools to automate the FM contouring and localization task.

6 Data Availability Statement

All code used is publicly available on GitHub. QSMxT for QSM reconstruction: https://qsmxt.github.io/ (v5.1.0). U-Net architecture and training code: https://github.com/astewartau/prostate.

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

7 Conflict of interest

The authors have no conflicts of interest to declare.

8 Acknowledgments

We thank the patients involved in this study and the radiation oncology team at Calvary Mater, Newcastle, Australia. The authors SB and AWS acknowledge funding through an ARC Linkage grant (LP200301393) and acknowledge support through the UQ AI Collaboratory. MB acknowledges funding from the Australian Research Council Future Fellowship grant FT140100865 and SR from the Austrian Science Fund (FWF): 31452. This research was partly conducted by the Australian Research Council Training Centre for Innovation in Biomedical Imaging Technology (project number IC170100035) and funded by the Australian Government. The authors would like to acknowledge Siemens Healthineers for supporting our project by providing the WIP sequence used to acquire the QSM data in this publication. This research was undertaken with the assistance of resources and services from the Queensland Cyber Infrastructure Foundation (QCIF) and the UQ Research Computing Centre (RCC).

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

[1] Richard J. Rebello, Christoph Oing, Karen E. Knudsen, et al. "Prostate cancer". In: *Nature Reviews Disease Primers* 7.1 (Feb. 2021), pp. 1–27. ISSN: 2056-676X. DOI: 10.1038/s41572-020-00243-0.

[2] Yaoqin Xie, David Djajaputra, Christopher R. King, et al. "Intrafractional Motion of the Prostate During Hypofractionated Radiotherapy". In: *International Journal of Radiation Oncology * Biology * Physics* 72.1 (Sept. 2008), pp. 236–246. ISSN: 0360-3016. DOI: 10.1016/j.ijrobp.2008.04.051.

[3] Simon J. Thomas, Mark Ashburner, George Samuel J. Tudor, et al. "Intra-fraction motion of the prostate during treatment with helical tomotherapy". In: *Radiotherapy and Oncology* 109.3 (Dec. 2013), pp. 482–486. ISSN: 0167-8140. DOI: 10.1016/j.radonc.2013.09.011.

[4] Michael Ng, Elizabeth Brown, Andrew Williams, et al. "Fiducial markers and spacers in prostate radiotherapy: current applications". In: *BJU International* 113.S2 (2014), pp. 13–20. ISSN: 1464-410X. DOI: 10.1111/bju.12624.

[5] John M. Schallenkamp, Michael G. Herman, Jon J. Kruse, et al. "Prostate position relative to pelvic bony anatomy based on intraprostatic gold markers and electronic portal imaging". In: *International Journal of Radiation Oncology * Biology * Physics* 63.3 (Nov. 2005), pp. 800–811. ISSN: 0360-3016. DOI: 10.1016/j.ijrobp.2005.02.022.

[6] Jarad Martin, Paul Keall, Shankar Siva, et al. "TROG 18.01 phase III randomised clinical trial of the Novel Integration of New prostate radiation schedules with adJuvant Androgen deprivation: NINJA study protocol". In: *BMJ Open* 9.8 (Aug. 2019). Publisher: British Medical Journal Publishing Group Section: Oncology, e030731. ISSN: 2044-6055, 2044-6055. DOI: 10.1136/bmjopen-2019-030731.

[7] Boris R. Tocco, Amar U. Kishan, Ting Martin Ma, et al. "MR-Guided Radiotherapy for Prostate Cancer". In: *Frontiers in Oncology* 10 (2020). ISSN: 2234-943X.

[8] Tony Young, Jason Dowling, Robba Rai, et al. "Clinical validation of MR imaging time reduction for substitute/synthetic CT generation for prostate MRI-only treatment planning". In: *Physical and Engineering Sciences in Medicine* 46.3 (Sept. 2023), pp. 1015–1021. ISSN: 2662-4737. DOI: 10.1007/s13246-023-01268-x.

[9] Jens M. Edmund and Tufve Nyholm. "A review of substitute CT generation for MRI-only radiation therapy". In: *Radiation Oncology* 12.1 (Jan. 2017), p. 28. ISSN: 1748-717X. DOI: 10.1186/s13014-016-0747-y.

[10] Jae Hyuk Choi, Danny Lee, Laura O'Connor, et al. "Bulk Anatomical Density Based Dose Calculation for Patient-Specific Quality Assurance of MRI-Only Prostate Radiotherapy". In: *Frontiers in Oncology* 9 (2019). ISSN: 2234-943X.

[11] David Bird, Michael G. Nix, Hazel McCallum, et al. "Multicentre, deep learning, synthetic-CT generation for ano-rectal MR-only radiotherapy treatment planning". In: *Radiotherapy and Oncology* 156 (Mar. 2021), pp. 23–28. ISSN: 0167-8140. DOI: 10.1016/j.radonc.2020.11.027.

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[12] Laura M. O'Connor, Jason A. Dowling, Jae Hyuk Choi, et al. "Validation of an MRI-only planning workflow for definitive pelvic radiotherapy". In: *Radiation Oncology* 17.1 (Mar. 2022), p. 55. ISSN: 1748-717X. DOI: 10.1186/s13014-022-02023-4.

[13] Joseph Weygand, Clifton David Fuller, Geoffrey S. Ibbott, et al. "Spatial Precision in Magnetic Resonance Imaging–Guided Radiation Therapy: The Role of Geometric Distortion". In: *International Journal of Radiation Oncology * Biology * Physics* 95.4 (July 2016), pp. 1304–1316. ISSN: 0360-3016. DOI: 10.1016/j.ijrobp.2016.02.059.

[14] J. N. H. Brunt. "Computed Tomography–Magnetic Resonance Image Registration in Radiotherapy Treatment Planning". In: *Clinical Oncology*. The Technical Aspects of Radiotherapy 22.8 (Oct. 2010), pp. 688–697. ISSN: 0936-6555. DOI: 10.1016/j.clon.2010.06.016.

[15] Soumya Ghose, Jhimli Mitra, David Rivest-Hénault, et al. "MRI-alone radiation therapy planning for prostate cancer: Automatic fiducial marker detection". In: *Medical Physics* 43.5 (2016), pp. 2218–2228. ISSN: 2473-4209. DOI: 10.1118/1.4944871.

[16] Matteo Maspero, Cornelis A. T. van den Berg, Frank Zijlstra, et al. "Evaluation of an automatic MR-based gold fiducial marker localisation method for MR-only prostate radiotherapy". In: *Physics in Medicine & Biology* 62.20 (Oct. 2017), p. 7981. ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa875f.

[17] Catarina Dinis Fernandes, Cuong V. Dinh, Marcel J. Steggerda, et al. "Prostate fiducial marker detection with the use of multi-parametric magnetic resonance imaging". In: *Physics and Imaging in Radiation Oncology* 1 (Jan. 2017), pp. 14–20. ISSN: 2405-6316. DOI: 10.1016/j.phro.2017.02.001.

[18] Christian Gustafsson, Juha Korhonen, Emilia Persson, et al. "Registration free automatic identification of gold fiducial markers in MRI target delineation images for prostate radiotherapy". In: *Medical Physics* 44.11 (2017), pp. 5563–5574. ISSN: 2473-4209. DOI: 10.1002/mp.12516.

[19] Christian Jamtheim Gustafsson, Johan Swärd, Stefan Ingi Adalbjörnsson, et al. "Development and evaluation of a deep learning based artificial intelligence for automatic identification of gold fiducial markers in an MRI-only prostate radiotherapy workflow". In: *Physics in Medicine & Biology* 65.22 (Nov. 2020), p. 225011. ISSN: 0031-9155. DOI: 10.1088/1361-6560/abb0f9.

[20] Ludovic de Rochefort, Ryan Brown, Martin R. Prince, et al. "Quantitative MR susceptibility mapping using piece-wise constant regularized inversion of the magnetic field". In: *Magnetic Resonance in Medicine* 60.4 (2008), pp. 1003–1009. ISSN: 1522-2594. DOI: 10.1002/mrm.21710.

[21] Karin Shmueli, Jacco A. de Zwart, Peter van Gelderen, et al. "Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data". In: *Magnetic Resonance in Medicine* 62.6 (2009), pp. 1510–1522. ISSN: 1522-2594. DOI: 10.1002/mrm.22135.

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[22] Ferdinand Schweser, Andreas Deistung, and Jürgen R. Reichenbach. "Foundations of MRI phase imaging and processing for Quantitative Susceptibility Mapping (QSM)". In: *Zeitschrift für Medizinische Physik* 26.1 (Mar. 2016), pp. 6–34. ISSN: 0939-3889. DOI: 10.1016/j.zemedi.2015.10.002.

[23] Andreas Deistung, Ferdinand Schweser, and Jürgen R. Reichenbach. "Overview of quantitative susceptibility mapping". In: *NMR in Biomedicine* 30.4 (2017), e3569. ISSN: 1099-1492. DOI: 10.1002/nbm.3569.

[24] Yi Wang, Pascal Spincemaille, Zhe Liu, et al. "Clinical quantitative susceptibility mapping (QSM): Biometal imaging and its emerging roles in patient care". In: *Journal of Magnetic Resonance Imaging* 46.4 (2017), pp. 951–971. ISSN: 1522-2586. DOI: 10.1002/jmri.25693.

[25] Thorsten A. Bley, Oliver Wieben, Christopher J. François, et al. "Fat and water magnetic resonance imaging". In: *Journal of Magnetic Resonance Imaging* 31.1 (2010), pp. 4–18. ISSN: 1522-2586. DOI: 10.1002/jmri.21895.

[26] Sina Straub, Julian Emmerich, Heinz-Peter Schlemmer, et al. "Mask-Adapted Background Field Removal for Artifact Reduction in Quantitative Susceptibility Mapping of the Prostate". In: *Tomography* 3.2 (June 2017), pp. 96–100. ISSN: 2379-139X. DOI: 10.18383/j.tom.2017.00005.

[27] Ryota Sato, Toru Shirai, Yoshihisa Soutome, et al. "Quantitative susceptibility mapping of prostate with separate calculations for water and fat regions for reducing shading artifacts". In: *Magnetic Resonance Imaging* 66 (Feb. 2020), pp. 22–29. ISSN: 0730-725X. DOI: 10.1016/j.mri.2019.11.006.

[28] Jannis Hanspach, Steffen Bollmann, Johanna Grigo, et al. "Deep learning–based quantitative susceptibility mapping (QSM) in the presence of fat using synthetically generated multi-echo phase training data". In: *Magnetic Resonance in Medicine* 88.4 (2022), pp. 1548–1560. ISSN: 1522-2594. DOI: 10.1002/mrm.29265.

[29] Beata Bachrata, Siegfried Trattnig, and Simon Daniel Robinson. "Quantitative susceptibility mapping of the head-and-neck using SMURF fat-water imaging with chemical shift and relaxation rate corrections". In: *Magnetic Resonance in Medicine* 87.3 (2022), pp. 1461–1479. ISSN: 1522-2594. DOI: 10.1002/mrm.29069.

[30] Ashley Wilton Stewart, Simon Daniel Robinson, Kieran O'Brien, et al. "QSMxT: Robust masking and artifact reduction for quantitative susceptibility mapping". In: *Magnetic Resonance in Medicine* 87.3 (2022), pp. 1289–1300. ISSN: 1522-2594. DOI: 10.1002/mrm.29048.

[31] Alexey V. Dimov, Jiahao Li, Thanh D. Nguyen, et al. "QSM Throughout the Body". In: *Journal of Magnetic Resonance Imaging* 57.6 (2023), pp. 1621–1640. ISSN: 1522-2586. DOI: 10.1002/jmri.28624.

[32] Sina Straub, Frederik B. Laun, Julian Emmerich, et al. "Potential of quantitative susceptibility mapping for detection of prostatic calcifications". In: *Journal of Magnetic Resonance Imaging* 45.3 (2017), pp. 889–898. ISSN: 1522-2586. DOI: 10.1002/jmri.25385.

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[33] Hirohito Kan, Takahiro Tsuchiya, Masato Yamada, et al. "Delineation of prostatic calcification using quantitative susceptibility mapping: Spatial accuracy for magnetic resonance-only radiotherapy planning". In: *Journal of Applied Clinical Medical Physics* 23.2 (2022), e13469. ISSN: 1526-9914. DOI: 10.1002/acm2.13469.

[34] Grace G. Zeng, Tom S. McGowan, Tessa M. Larsen, et al. "Calcifications Are Potential Surrogates for Prostate Localization in Image-Guided Radiotherapy". In: *International Journal of Radiation Oncology * Biology * Physics* 72.4 (Nov. 2008), pp. 963–966. ISSN: 0360-3016. DOI: 10.1016/j.ijrobp.2008.07.021.

[35] Reyhaneh Nosrati, Moti Paudel, Ananth Ravi, et al. "Potential applications of the quantitative susceptibility mapping (QSM) in MR-guided radiation therapy". In: *Physics in Medicine & Biology* 64.14 (July 2019), p. 145013. ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab2623.

[36] Reyhaneh Nosrati, Abraam Soliman, Habib Safigholi, et al. "MRI-based automated detection of implanted low dose rate (LDR) brachytherapy seeds using quantitative susceptibility mapping (QSM) and unsupervised machine learning (ML)". In: *Radiotherapy and Oncology* 129.3 (Dec. 2018), pp. 540–547. ISSN: 0167-8140. DOI: 10.1016/j.radonc.2018.09.003.

[37] Reyhaneh Nosrati, William Y. Song, Matthew Wronski, et al. "Feasibility of an MRI-only workflow for postimplant dosimetry of low-dose-rate prostate brachytherapy: Transition from phantoms to patients". In: *Brachytherapy* 18.6 (Nov. 2019), pp. 863–874. ISSN: 1538-4721. DOI: 10.1016/j.brachy.2019.06.004.

[38] R. Nosrati, W. W. Lam, M. Paudel, et al. "Feasibility of using a single MRI acquisition for fiducial marker localization and synthetic CT generation towards MRI-only prostate radiation therapy treatment planning". In: *Biomedical Physics & Engineering Express* 7.4 (June 2021), p. 045016. ISSN: 2057-1976. DOI: 10.1088/2057-1976/ac0501.

[39] Vladimír Jellúš and Stephan AR Kannengiesser. "Adaptive coil combination using a body coil scan as phase reference". In: *ISMRM 22nd Annual Meeting & Exhibition, Milan, Italy*. Abstract #4406. International Society for Magnetic Resonance in Medicine, 2014.

[40] Ashley Wilton Stewart. *QSMxT: A Complete QSM Processing Framework*. original-date: 2020-12-21T00:24:36Z. 2023.

[41] Christian Langkammer, Kristian Bredies, Benedikt A. Poser, et al. "Fast quantitative susceptibility mapping using 3D EPI and total generalized variation". In: *NeuroImage* 111 (May 2015), pp. 622–630. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2015.02.041.

[42] Christian Langkammer, Kristian Bredies, Benedikt A. Poser, et al. *TGV QSM Algorithm: Python Source Code*. Publication Title: Neuroimaging. May 2015.

[43] Korbinian Eckstein, Beata Bachrata, Gilbert Hangel, et al. "Improved susceptibility weighted imaging at ultra-high field using bipolar multi-echo acquisition and optimized image processing: CLEAR-SWI". In: *NeuroImage* 237 (Aug. 2021), p. 118175. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2021.118175.

[44] G.e. Hagberg, I. Indovina, J.n. Sanes, et al. "Real-time quantification of T changes using multiecho planar imaging and numerical methods". In: *Magnetic Resonance in Medicine* 48.5 (2002), pp. 877–882. ISSN: 1522-2594. DOI: 10.1002/mrm.10283.

bioRxiv preprint doi: https://doi.org/10.1101/2023.10.26.564293; this version posted October 31, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[45] Korbinian Eckstein. *MriResearchTools*. korbinian90/MriResearchTools.jl. 2023.

[46] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, et al. "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability". In: *NeuroImage* 31.3 (July 2006), pp. 1116–1128. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2006.01.015.

[47] Sathiesh Kaliyugarasan and Alexander Lundervold. *MMIV-ML/fastMONAI: Simplifying deep learning for medical imaging*. fastmonai.io. Feb. 2023.

[48] Satheshkumar Kaliyugarasan and Alexander S. Lundervold. "fastMONAI: A low-code deep learning library for medical image analysis". In: *Software Impacts* 18 (Nov. 2023), p. 100583. ISSN: 2665-9638. DOI: 10.1016/j.simpa.2023.100583.

[49] Jeremy Howard and Sylvain Gugger. "Fastai: A Layered API for Deep Learning". In: *Information* 11.2 (Feb. 2020), p. 108. ISSN: 2078-2489. DOI: 10.3390/info11020108.

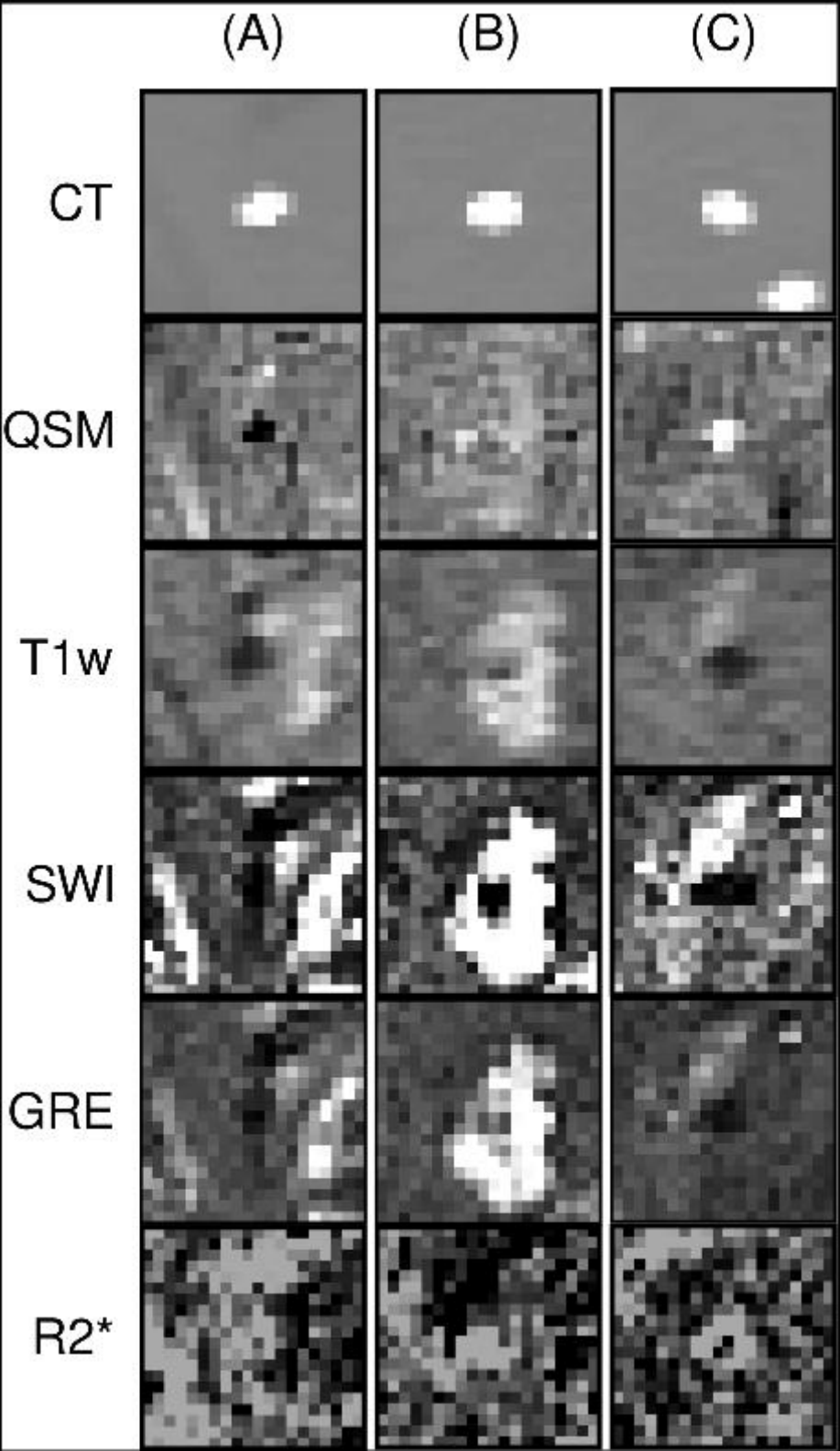
[50] M. Jorge Cardoso, Wenqi Li, Richard Brown, et al. *MONAI: An open-source framework for deep learning in healthcare*. Nov. 2022. DOI: 10.48550/arXiv.2211.02701.

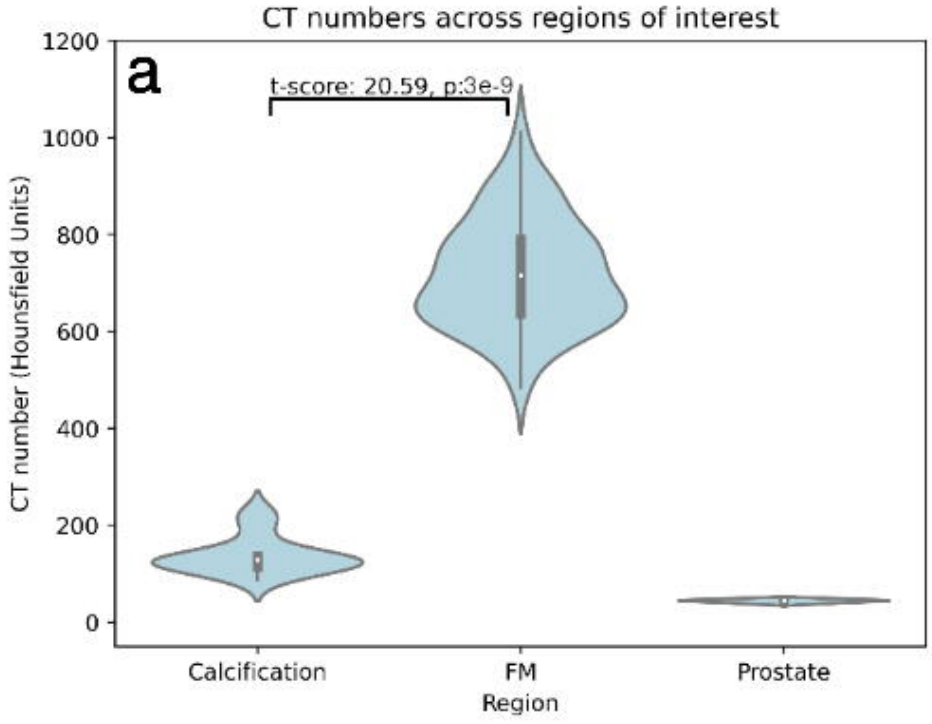
[51] Sagar Imambi, Kolla Bhanu Prakash, and G. R. Kanagachidambaresan. "PyTorch". In: *Programming with TensorFlow: Solution for Edge Computing Applications*. Ed. by Kolla Bhanu Prakash and G. R. Kanagachidambaresan. EAI/Springer Innovations in Communication and Computing. Cham: Springer International Publishing, 2021, pp. 87–104. ISBN: 978-3-030-57077-4. DOI: 10.1007/978-3-030-57077-4_10.

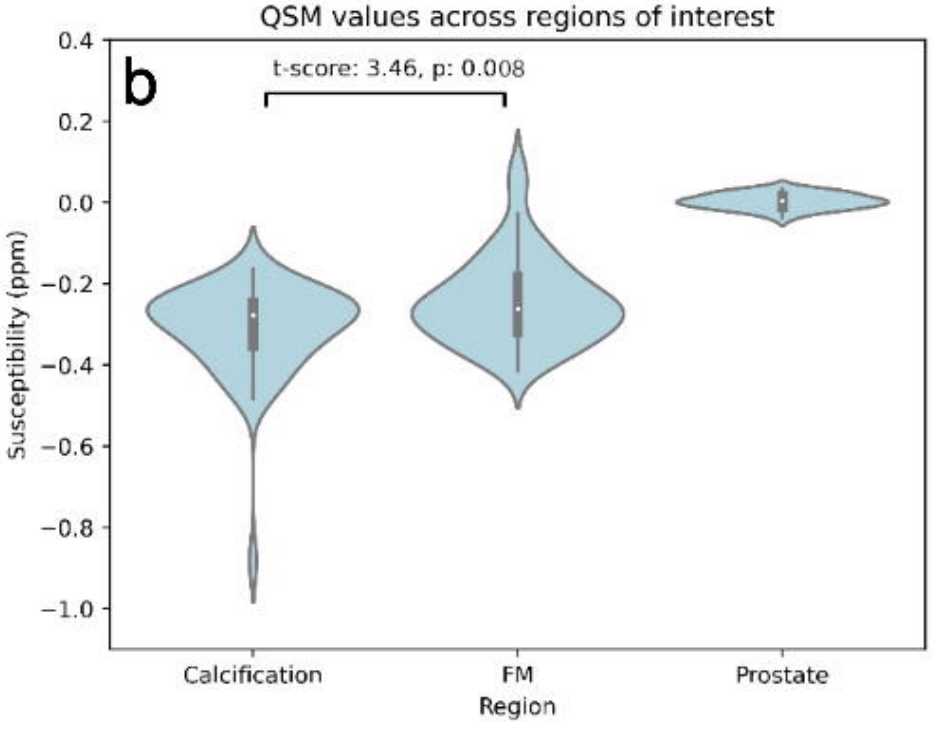
[52] Berkin Bilgic, Mauro Costagli, Kwok-Shing Chan, et al. *Recommended Implementation of Quantitative Susceptibility Mapping for Clinical Research in The Brain: A Consensus of the ISMRM Electro-Magnetic Tissue Properties Study Group*. July 2023. DOI: 10.48550/arXiv.2307.02306.

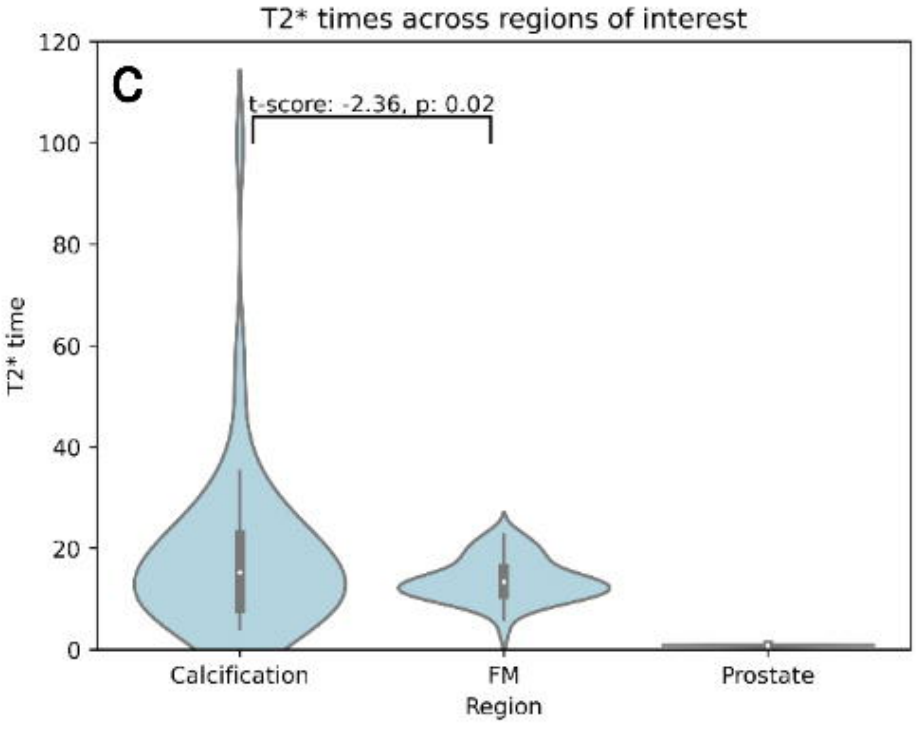
[53] Christos G. Xanthis, Dimitrios Filos, Kostas Haris, et al. "Simulator-generated training datasets as an alternative to using patient data for machine learning: An example in myocardial segmentation with MRI". In: *Computer Methods and Programs in Biomedicine* 198 (Jan. 2021), p. 105817. ISSN: 0169-2607. DOI: 10.1016/j.cmpb.2020.105817.

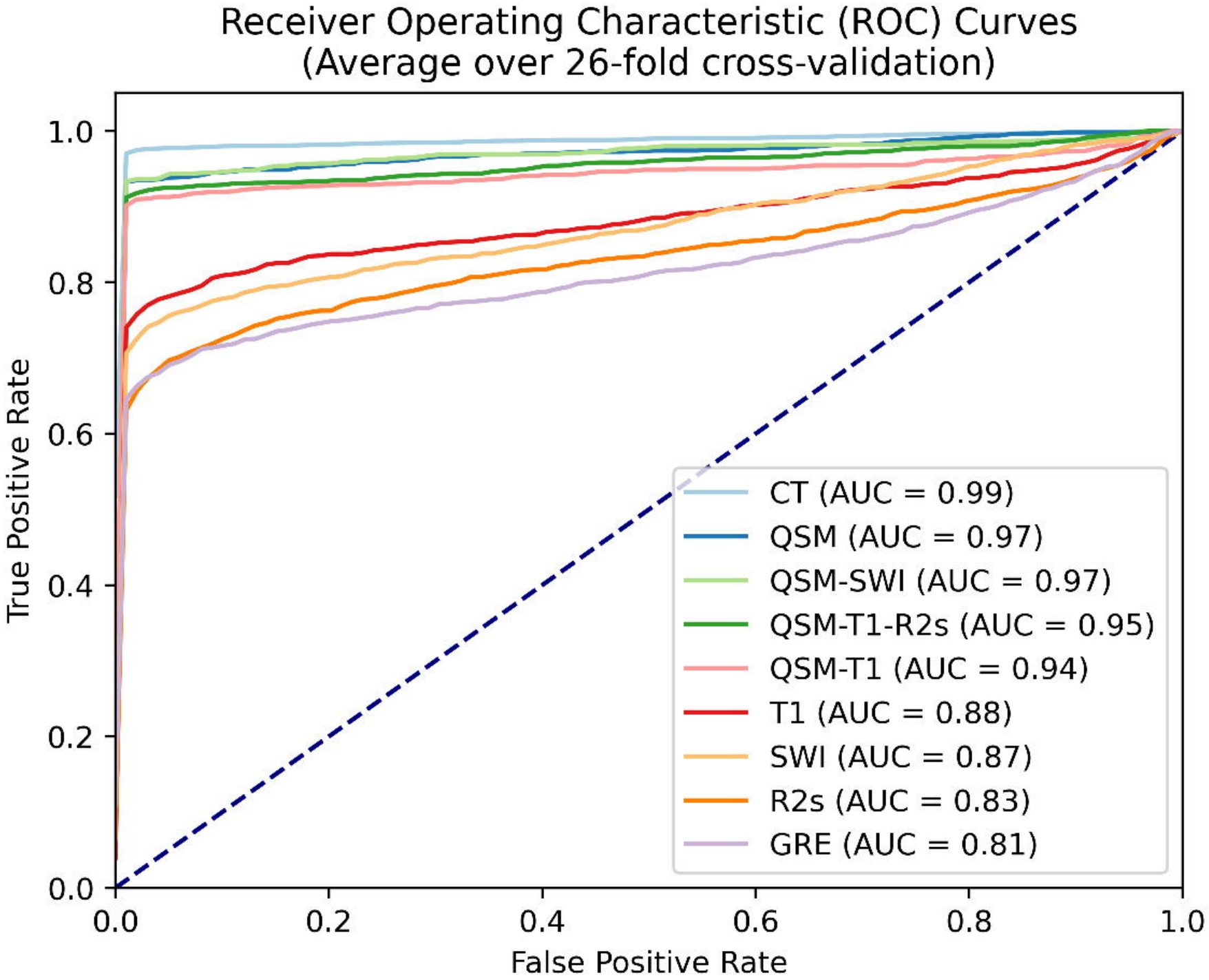
[54] K. Sepehri, X. Song, R. Proulx, et al. "Towards effective machine learning in medical imaging analysis: A novel approach and expert evaluation of high-grade glioma 'ground truth' simulation on MRI". In: *International Journal of Medical Informatics* 146 (Feb. 2021), p. 104348. ISSN: 1386-5056. DOI: 10.1016/j.ijmedinf.2020.104348.

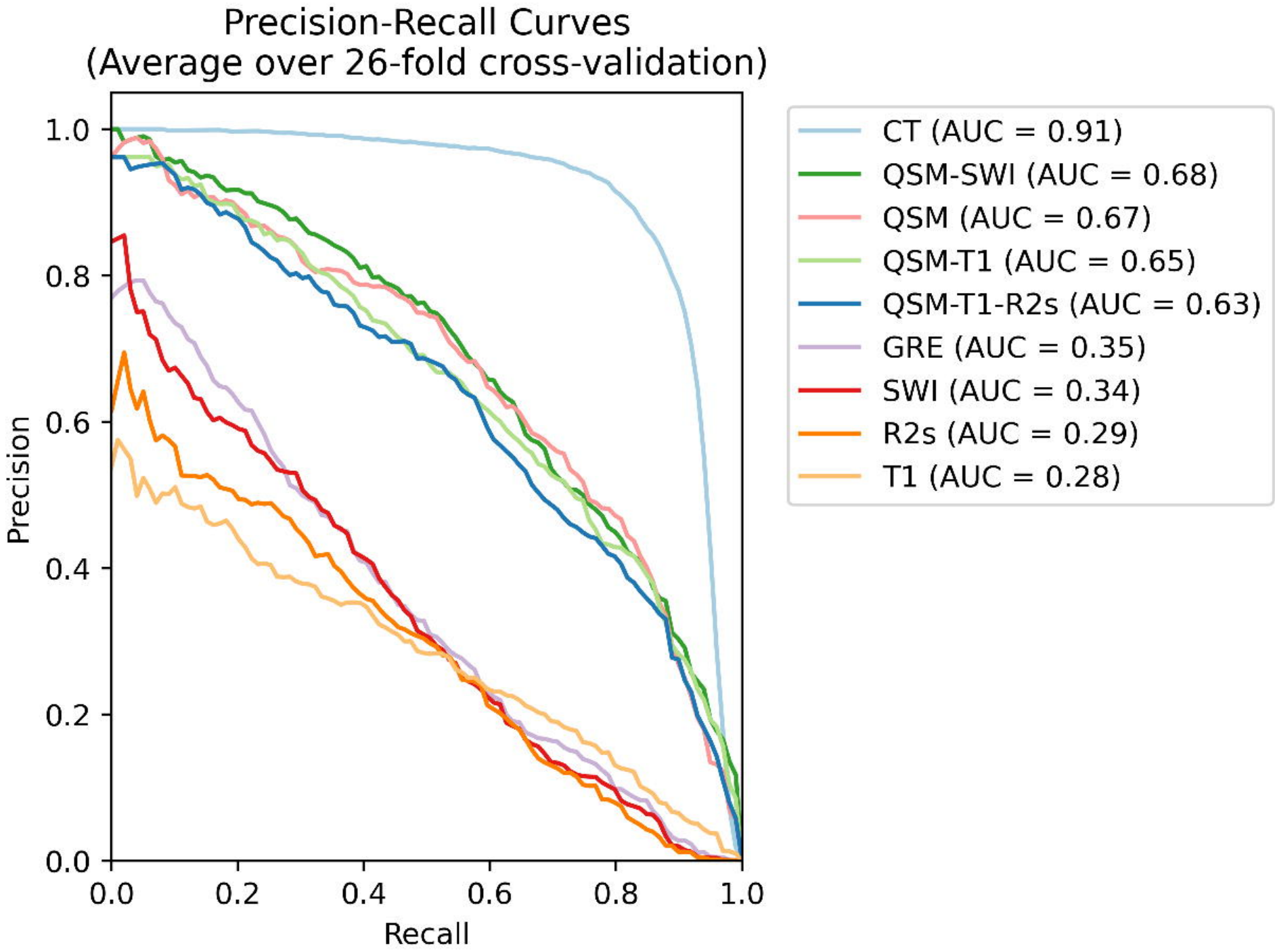

(A)
(B)
(C)
CT
QSM
T1w
SWI
GRE
R2*

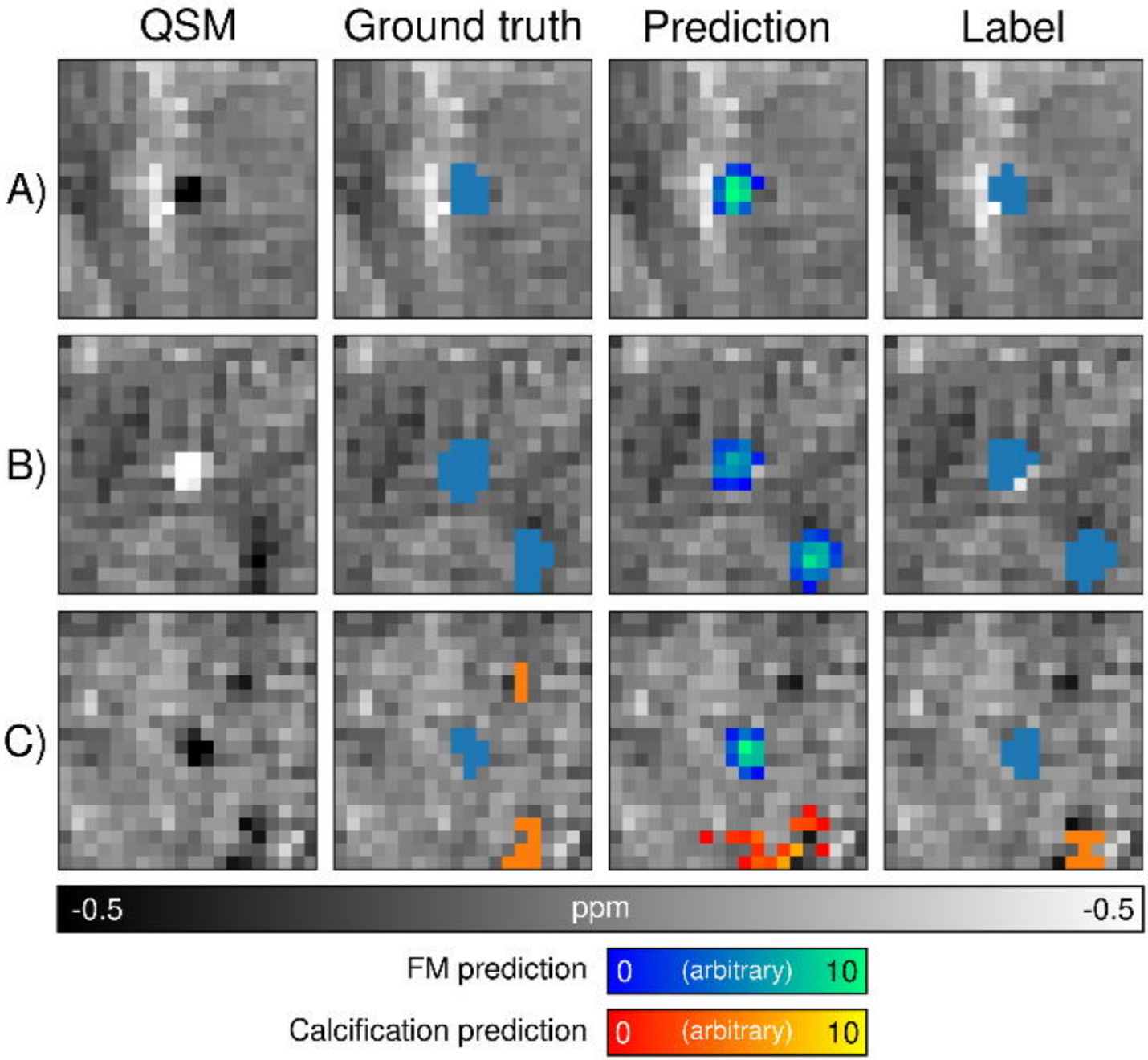

CT numbers across regions of interest
a
t-score: 20.59, p:3e-9
CT number (Hounsfield Units)
1200
1000
800
600
400
200
0
Calcification
FM
Prostate
Region


QSM values across regions of interest
b
t-score: 3.46, p: 0.008
Susceptibility (ppm)
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
Calcification
FM
Prostate
Region


T2* times across regions of interest
c
t-score: -2.36, p: 0.02
T2* time
120
100
80
60
40
20
0
Calcification
FM
Prostate
Region


Receiver Operating Characteristic (ROC) Curves
(Average over 26-fold cross-validation)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
CT (AUC = 0.99)
QSM (AUC = 0.97)
QSM-SWI (AUC = 0.97)
QSM-T1-R2s (AUC = 0.95)
QSM-T1 (AUC = 0.94)
T1 (AUC = 0.88)
SWI (AUC = 0.87)
R2s (AUC = 0.83)
GRE (AUC = 0.81)


Precision-Recall Curves
(Average over 26-fold cross-validation)
Precision
Recall
0.0
0.2
0.4
0.6
0.8
1.0
CT (AUC = 0.91)
QSM-SWI (AUC = 0.68)
QSM (AUC = 0.67)
QSM-T1 (AUC = 0.65)
QSM-T1-R2s (AUC = 0.63)
GRE (AUC = 0.35)
SWI (AUC = 0.34)
R2s (AUC = 0.29)
T1 (AUC = 0.28)


QSM
Ground truth
Prediction
Label
A)
B)
C)
-0.5
ppm
-0.5
FM prediction
0
(arbitrary)
10
Calcification prediction
0
(arbitrary)
10