MRI-Based Radiomics as a Mechanism for Bone Health Assessment: An Exploratory Feasibility Study

Kevin M. Koch[1], Brad Swearningen[1], Ryan Breighner, Matthew F. Koff[2], Timothy M. Wright[2], Fernando Quevedo-Gonzalez[2], and Hollis G. Potter[2]

1. Medical College of Wisconsin, Milwaukee, WI
2. Hospital for Special Surgery, New York, NY

kmkoch@mcw.edu

Disclosures: None

Introduction: Successful implant osteointegration after total knee replacement (TKR) provides stability and longevity to the replacement. However, conventional imaging modalities for post-operative bone assessment, such as computed tomography (CT), are compromised by metal artifacts introduced by the implant components. These artifacts can obscure regions of interest adjacent to the device, thereby constraining a comprehensive assessment of osteointegration. Magnetic resonance imaging (MRI), with its innate advantage in soft tissue contrast and the absence of ionizing radiation, offers an attractive alternative for osteointegration assessment. Recent advancements in metal-artifact reduction sequences further augment MRI's utility. Nonetheless, leveraging MRI to evaluate osteointegration in the presence of metallic implants has yet to be attempted.

Modeling of bone health using MRI-based radiomics, an emerging discipline focused on extracting high-dimensional image features, presents a promising solution to this technical challenge. As a feasibility study of this concept, the present analysis focused on pre-operative MRI and CT datasets for which clear comparisons in the absence of metal artifacts can be performed. With an analysis focused on regions of interest relevant to TKR in the proximal tibia and distal femur, and utilizing the sophisticated auto-machine learning tools provided by the h2o library, the objective of this study was to establish the feasibility of MRI radiomic analysis as a viable investigational approach for post-operative osteointegration analysis after TKR.

Methods: *Study Cohort:* A total of 14 subjects, consisting of 7 males (mean age: 51.2 ± 21.1 years) and 7 females (mean age: 57.0 ± 20.2 years), who had undergone pre-surgical clinical CT and MRI exams, were selected. Of these, 12 subjects were pre-screened for TKR, while the remaining two were evaluated for ACL reconstruction. All data were collected within the context of an approved IRB protocol. *MRI and CT:* CT imaging was acquired using standard clinical protocols, specific to clinical concern (120-140 kVp and 200-330 mA). 2D multi-slice fast-spin-echo based MRI data were collected with two specific contrasts: fat-suppressed inversion recovery (IR) and proton-density (PD). Images were acquired in three orthogonal scan planes using established clinical protocols and standard patient positioning [1]. *Image Preprocessing:* From the CT scans, the proximal femur and tibia bones were auto-segmented. These segments served as references to which MRI volumes were independently registered. Subsequently, square patches from the MRI and CT pairs were extracted from these regions. Four patch sizes, with side lengths of 8, 16, 32, and 64 voxels, representing respective areas of (0.16, 0.65, 2.56, 10.24) cm^2, were considered for modeling purposes. *Radiomics & Bone Density Estimation:* Using Pyradiomics [2], first and second-order radiomic features were extracted from each patch, totaling 96 features per patch. Bone density approximation was done using the mean HU values within the segmented region of the CT patches. Patches presenting a mean HU value > 1000 were excluded due to conventional MRI's inherent limitation in imaging cortical bone. *Model Computation and Analysis:* Associative models between MR radiomic features and CT bone density estimates were constructed using the h2o auto-machine learning platform [3], using 25 potential machine-learning modeling approaches. This infrastructure was also utilized to assess model performance using metrics of root-mean square error (RMSE), mean adjusted error (MAE), and R^2 values. Patch-wise data for each model were split into 70% training, 15% validation, and 15% testing.

Results: Representative MRI and CT images, along with example analysis patches, are presented in Figure 1. Consolidated model performance metrics, as illustrated in Figure 2, indicate a clear distinction between the two analyzed MRI contrasts (PD and IR). On average, models trained on the PD contrast outperformed those trained on the IR contrast across all evaluated metrics. Specifically, the PD contrast led to lower (RMSE) and (MAE) values, and yielded higher R^2 values across the testing patches.

Detailed examination of the top-performing models, with performance metrics summarized in Table 1 for tibia and femur bone regions, revealed that Gradient Boosting Machine approach (GBM) consistently offered top modeling performance . For the femur (with IR contrast in the Axial plane), a GBM model provided an R^2 of 0.73, while a GBM model also provided the best performance in the tibia using PD contrast in the axial plane. Other notable high performing models in the h2o leaderboard include Distributed Random Forest (DRF) and Extremely Randomized Trees (XRT), which exhibited testing errors only 2% higher than the top performing model.

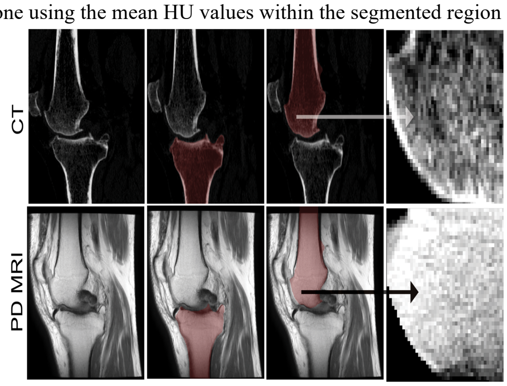

Figure 1: Sample CT and MRI images of a segmented femur, displaying a 64x64 analysis patch used for texture feature extraction.

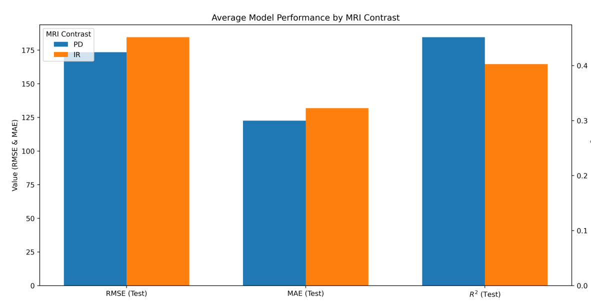

Figure 2: Performance across all models as a function of MRI contrast

Parameter	Best Result (Tibia)	Best Result (Femur)
Bone	Tibia	Femur
R^2	0.69	0.74
MRI Contrast	PD	IR
Plane	AX	AX
Patch Size (X)	64	64
Number of Patches	15874	10807

Table 1: Performance of best models for each key ROI (bone region)

Discussion: Our findings highlight the potential of using radiomic features from MRI to assess osseointegration post TKR, a key indicator of long-term success. The nuanced modeling performances enabled by the h2o auto-machine learning framework has emphasized the potential of GBM as an optimal modeling infrastructure. Interestingly, while models utilizing the PD contrast generally showcased better performance, the top-performing model across the survey was derived from IR MR contrast. This observation may hint at the presence of certain radiomic signatures in the IR contrast that are particularly pertinent to bone-health assessment. The general efficacy of the PD contrast, juxtaposed with the standout performance of a single IR-based model, underscores the complexity of MRI-based osseointegration evaluation. Future studies could further probe this contrast-based discrepancy using larger datasets, potentially unveiling deeper insights into post-operative bone health and implant integration. Lastly, the utilization of a bone-density calibration phantom in prospective studies will increase the accuracy of the model in predicting bone density.

Clinical Significance: Leveraging MRI radiomic features to evaluate osseointegration after total knee replacement offers clinicians a sophisticated, non-invasive tool to monitor post-operative success and implant stability. This advancement holds the potential to transform post-operative care, allowing for timely interventions if poor osseointegration is detected.

References: [1] *Eagle, S et al. (2017). J. Ortho. Res, 35(3), 412-423.* [2] *van Griethuysen, et al. (2017). Cancer Research, 77(21), e104–e107.*, [3] https://h2o.ai