Combining multiparametric MRI with receptor information to optimize prediction of pathologic response to neoadjuvant therapy in breast cancer: preliminary results

Hakmook Kang
Allison Hainline
Lori R. Arlinghaus
Stephanie Elderidge
Xia Li
Vandana G. Abramson
Anuradha Bapsi Chakravarthy
Richard G. Abramson
Brian Bingham
Kareem Fakhoury
Thomas E. Yankeelov

Hakmook Kang, Allison Hainline, Lori R. Arlinghaus, Stephanie Elderidge, Xia Li, Vandana G. Abramson, Anuradha Bapsi Chakravarthy, Richard G. Abramson, Brian Bingham, Kareem Fakhoury, Thomas E. Yankeelov, "Combining multiparametric MRI with receptor information to optimize prediction of pathologic response to neoadjuvant therapy in breast cancer: preliminary results," *J. Med. Imag.* 5(1), 011015 (2017), doi: 10.1117/1.JMI.5.1.011015.

Combining multiparametric MRI with receptor information to optimize prediction of pathologic response to neoadjuvant therapy in breast cancer: preliminary results

Hakmook Kang,^{a,b,*} Allison Hainline,^a Lori R. Arlinghaus,^c Stephanie Elderidge,^{d,e} Xia Li,^f Vandana G. Abramson,^{g,h} Anuradha Bapsi Chakravarthy,^{g,i} Richard G. Abramson,^{b,j} Brian Bingham,^k Kareem Fakhoury,^k and Thomas E. Yankeelov^{d,e}

^aVanderbilt University Medical Center, Department of Biostatistics, Nashville, Tennessee, United States

^bVanderbilt University Medical Center, Center for Quantitative Sciences, Nashville, Tennessee, United States

^cVanderbilt University Medical Center, Institute of Imaging Science, Nashville, Tennessee, United States

^dUniversity of Texas, Institute of Computational and Engineering Sciences, Austin, Texas, United States

^eUniversity of Texas, Department of Biomedical Engineering, Austin, Texas, United States

^fGE Global Research, Niskayuna, New York, United States

^gVanderbilt University Medical Center, Ingram Cancer Center, Nashville, Tennessee, United States

^hVanderbilt University Medical Center, Medical Oncology, Nashville, Tennessee, United States

ⁱVanderbilt University Medical Center, Department of Radiation Oncology, Nashville, Tennessee, United States

^jVanderbilt University Medical Center, Department of Radiology and Radiological Science, Nashville, Tennessee, United States

^kVanderbilt University Medical Center, School of Medicine, Nashville, Tennessee, United States

Abstract Pathologic complete response following neoadjuvant therapy (NAT) is used as a short-term surrogate marker of eventual outcome in patients with breast cancer. Analyzing voxel-level heterogeneity in MRI-derived parametric maps, obtained before and after the first cycle of NAT ($n = 33$), in conjunction with receptor status, may improve the predictive accuracy of tumor response to NAT. Toward that end, we incorporated two MRI-derived parameters, the apparent diffusion coefficient and efflux rate constant, with receptor status in a logistic ridge-regression model. The area under the curve (AUC) and Brier score of the model computed via 10-fold cross validation were 0.94 (95% CI: 0.85, 0.99) and 0.11 (95% CI: 0.06, 0.16), respectively. These two statistics strongly support the hypothesis that our proposed model outperforms the other models that we investigated (namely, models without either receptor information or voxel-level information). The contribution of the receptor information was manifested by an 8% to 15% increase in AUC and a 14% to 21% decrease in Brier score. These data indicate that combining multiparametric MRI with hormone receptor status has a high likelihood of improved prediction of pathologic response to NAT in breast cancer. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.5.1.011015]

Keywords: spatial heterogeneity; imaging biomarkers of response; quantitative imaging; predictive value of tests; dynamic contrast-enhanced-magnetic resonance imaging; diffusion-weighted-magnetic resonance imaging.

Paper 17192SSRR received Jun. 30, 2017; accepted for publication Dec. 5, 2017; published online Dec. 29, 2017.

1 Introduction

Neoadjuvant therapy (NAT) not only provides a way to decrease tumor size to allow for breast conservation therapy but also is an excellent opportunity to observe tumor sensitivity to a particular regimen. If the primary tumor responds, any systemic micrometastases may also be sensitive to the same therapy.¹ Conversely, if the primary tumor is unresponsive, then the treatment can be changed to another, potentially more effective approach. A number of preoperative systemic regimens are available, some targeting specific cancer subtype according to the National Comprehensive Cancer Network (NCCN). For example, in patients with human epidermal growth factor receptor 2 (HER2) negative disease, these include dose-dense AC (doxorubicin/cyclophosphamide) and dose-dense AC followed by weekly paclitaxel or TC (docetaxel/cyclophosphamide). In patients with

HER2 positive disease, AC followed by TH +/- *P* (doxorubicin/cyclophosphamide followed by paclitaxel and trastuzumab with or without pertuzumab) or TCH +/- *P* (docetaxel plus carboplatin/trastuzumab with or without pertuzumab) are available. Numerous other regimens are also available (see NCCN guidelines for a more complete listing, Ref. 2). The development of a method used to predict response early in the course of NAT would allow individuals to receive a treatment tailored to their specific agents. Unfortunately, evaluating the effectiveness of NAT currently requires a period of 3 to 5 months as described in NCCN guidelines, during which time one may see the potential use of ineffective and harmful agents. We propose that integrating quantitative magnetic resonance imaging (MRI) data with hormone receptor status can enable the accurate prediction of a treatment response after an individual patient has received a single cycle of therapy.

*Address all correspondence to: Hakmook Kang, E-mail: hakmook.kang@vanderbilt.edu

The response of breast tumors to NAT is currently evaluated by large changes in tumor size as measured by a physical exam, conventional MRI, and/or ultrasound, but these methods generally do not show whether a tumor is responding until the patient has received several treatment cycles. To overcome this hurdle, many investigators have applied diffusion-weighted MRI (DW-MRI) and dynamic contrast-enhanced MRI (DCE-MRI) to provide quantitative and complementary information in monitoring the response of breast tumors to NAT.^{3–17}

DW-MRI is sensitive to the thermally induced, random motion of water molecules in tissue. The rate of diffusion in tissue is described by an apparent diffusion coefficient (ADC), which can be measured by established DW-MRI methods.^{18–21} In controlled experiments, the variations in ADC have been shown to correlate inversely with tissue cellularity.^{22–26} In DCE-MRI, images are acquired before and after a contrast agent is injected into a peripheral vein. Each image in this dynamic series corresponds to one time point, and temporal intensity variation for each voxel can then be analyzed with a pharmacokinetic model to estimate physiological parameters such as the contrast agent transfer constant (K^{trans} related to vessel perfusion and permeability), the extravascular extracellular volume fraction (v_e), the plasma volume (v_p), and the efflux rate constant ($k_{\text{ep}} = K^{\text{trans}}/v_e$).^{27–34}

An emerging area of interest is predicting response to NAT using averaged quantitative, MRI-derived parameters over a region of interest (ROI) in combination with clinical variables.³⁵ Importantly, other studies have demonstrated that combining voxel-level DCE- and DW-MRI³⁶ and ROI-averaged multimodal MR and/or PET/CT^{7,9} metrics (e.g., k_{ep} from DCE-MRI, ADC from DW-MRI, and maximum standardized uptake value from PET/CT) enhances the predictive accuracy of the response to the first cycle of NAT, compared with that of each modality alone. Moreover, it has been demonstrated that breast cancer subtypes can significantly affect the prediction of tumor response to NAT.^{37–40} This prompted research to include breast cancer subtype, along with the clinical variables and voxel-level quantitative parameters derived from DCE- and DW-MRI as the major predictors in the NAT response model.

During the past 15 years, breast cancer classifications have evolved from those based entirely on histopathology to those based on gene expression profiling.⁴¹ In practice, cancer subtypes can be determined either directly via multigene assays or indirectly using immunochemistry. The following subtypes are usually defined: luminal A [estrogen receptor (ER) and/or progesterone receptor (PR) positive, HER2 negative], luminal B (ER and/or PR positive, HER2 positive), HER2 (ER negative, PR negative, and HER2 positive), and basal like (ER negative, PR negative, HER2 negative, i.e., triple negative).³⁸ For luminal B and HER2 molecular subtype breast cancers, it has been shown that preoperative MRI can help identify multifocal and/or multicentric disease and lymph node involvement.⁴² Given that the selection of treatment for breast cancer patients could be guided by molecular subtypes, we hypothesized that including information on the molecular subtypes would improve the predictive ability of our model.

Herein, we proposed a data-fusion approach via a logistic ridge-regression model, which could enhance predictive ability of MRI-derived quantitative variables for early detection of a breast cancer patient's response to NAT using the patient's hormone receptor information along with clinical information. For baseline comparison, we also employed a more conventional

predictive model based on response evaluation criteria in solid tumors (RECIST) and a previously established model that did not include information on receptor status.³⁶ We assessed the predictive power of each model in terms of the area under the receiver operator curves (i.e., AUCs) and Brier scores computed via 10-fold cross validation (CV).

2 Methods

Our previous study³⁶ demonstrated that including both voxel-level k_{ep} and ADC in a prediction model could enhance predictive ability of tumor response to NAT by employing redundancy analysis and logistic ridge regression. The same approach was employed to rigorously assess the contribution of hormone receptor information in prediction models via 10-fold CV in this work. Moreover, the contribution of both voxel-level k_{ep} and ADC was assessed, compared with RECIST and clinical information, which was not addressed in the previous study.

2.1 Patient Population and Magnetic Resonance Imaging Data

All patients signed consent forms prior to enrollment. This study was approved by Vanderbilt University Institutional Review Board. Breast cancer patients ($n = 33$) undergoing neoadjuvant chemotherapy (see Table 1 in Ref. 36 for a list of neoadjuvant treatment regimens and receptor status) as a component of care were recruited and scanned for DW- and DCE-MRI on a Philips 3.0T Achieva MRI scanner (Philips Healthcare, Best, The Netherlands) at the following three time points: before NAT (t_1), after the first cycle of NAT (t_2), and after all cycles of NAT (t_3). The median duration between t_1 and t_2 was 14 days [Q_1 (first quartile) = 12 days and Q_3 (third quartile) = 18.5 days] and the median total number of cycles of treatment was 14.5 cycles ($Q_1 = 12$ cycles and $Q_3 = 16$ cycles). The median duration of entire NAT was 103 days ($Q_1 = 84$ days and $Q_3 = 133$ days). Following the completion of NAT, pathologic complete response (pCR) and nonresponse (non-pCR) were determined by a breast pathologist at the time of surgery. In this study population, 12 patients exhibited pCR and 21 patients achieved non-pCR.

All patients had histologically documented invasive breast cancer at least 1 cm in the longest dimension with a sufficient risk of recurrence, to warrant the use of NAT. This risk was determined by the treating oncologist using pretreatment pathologic characteristics including tumor size, nodal status, grade, ER, PR, Ki-67, and HER2 status as measured by immunohistochemistry test on a patient's biopsy sample. In particular, HER2 was scored as positive with the immunohistochemical staining score of 3+ or 2+ and the amplification ratio of

Table 1 Summary of classification of molecular subtypes of breast cancer by receptor status.

Subtype	Receptor status	N (%)
Luminal A	ER+, and/or PR+, HER2–	9 (25)
Luminal B	ER+, and/or PR+, HER2+	6 (18)
HER2	ER–, PR–, and HER2+	8 (24)
Basal	ER–, PR–, and HER2–	10 (31)

≥ 2.2 on fluorescence *in situ* hybridization.⁴³ Estrogen and PR status was scored as positive when at least 1% of tumor cells showed positive nuclear staining of any intensity. Receptor status was considered negative if <1% of tumor cells showed staining of any intensity.⁴⁴

A full description of patients, MRI data acquisition, image analysis methods to derive k_{ep} and ADC, and the registration of longitudinally acquired parametric maps can be found in Ref. 36. Note that the registration of longitudinally acquired parametric maps^{45–48} was crucial for obtaining changes at the voxel-level for ADC and k_{ep} .

2.2 Statistical Analysis

As described in Ref. 36, ADC and k_{ep} data for each voxel in each tumor slice were acquired from 33 patients before (t_1) and after (t_2) the first cycle of NAT. These data were combined with tumor grade and age information to predict tumor response at the completion of NAT. Patients were identified as pCR if there was an absence of viable tumor cells in the primary tumor bed or lymph nodes at the time of surgery. Otherwise, the patient was identified as a nonresponder. For both k_{ep} and ADC, the subsets of voxels that displayed an increase in k_{ep} and a decrease in ADC from t_1 to t_2 were subjected to a redundancy analysis⁴⁹ to select a set of nonpredictable variables (defined in Appendix A). The redundancy analysis was employed after dividing the voxel-based parameter values (e.g., ADC values in the subset defined in the previous sentence) into 20 equal bins (i.e., min, 5th, . . . , 95th, max). The percentile values selected via the redundancy analysis (please see Appendix A for details) were employed as explanatory variables in subsequent logistic ridge-regression models. Due to the small number of patients ($n = 33$) compared with the number of explanatory variables in logistic regression models, we employed logistic ridge-regression models to avoid overfitting.⁵⁰

In addition to the MRI data, we used binary hormone receptor status measured at t_1 as covariates in the prediction models; i.e., ER = 0 or 1, PR = 0 or 1, and HER2 = 0 or 1, where 0 indicates a negative and 1 indicates a positive expression. In each model, pCR (0 = nonresponder and 1 = responder) serves as the binary outcome variable, whereas the selected percentiles of k_{ep} and ADC and hormonal receptor status information were explanatory variables in conjunction with tumor grade and patient age. We also added the mean of each parameter within an ROI as explanatory variables into the model, which would be major explanatory variables when applied voxel-level information is not available. Therefore, the voxel-clinical-receptor (VCR) model is expressed as follows:

$$\begin{aligned} \text{logit}[Pr(pCR = 1)] = & \beta_0 + \beta_1 h_{kep1} + \beta_2 h_{kep2} + \cdots \\ & + \beta_m h_{ADC1} + \cdots + \beta_N h_{ADCN} \\ & + \beta_{N+1} M_{kep} + \beta_{N+2} M_{ADC} \\ & + \beta_{N+3} ER + \beta_{N+4} PR + \beta_{N+5} HER2 \\ & + \beta_{N+6} Age + \beta_{N+7} Tumor, \end{aligned} \quad (1)$$

where h_{kepi} and h_{ADCi} are the i 'th percentiles for k_{ep} and ADC, respectively, M_{kep} and M_{ADC} denote the ROI-mean of k_{ep} and ADC, respectively, ER, PR, and HER2 denote the corresponding hormone receptor status (binary), age and tumor denote the

patient age (continuous) and tumor grade (continuous), and β 's are regression coefficients.

For comparisons, five other models were analyzed including (1) a model with a change in the longest dimension of the tumor (RECIST) from t_1 to t_2 , hormone receptor status, age, and tumor [RECIST-clinical-receptor (RCR) model], (2) the RCR model with RECIST removed [clinical-receptor (CR) model], (3) the VCR model with ER, PR, and HER2 data removed [voxel-clinical information (VC) model], (4) the RCR model with the receptor information removed [RECIST-clinical (RC) model], and (5) the CR model with the receptor information removed [clinical-only (CO) model]. In each logistic ridge regression, the tuning parameter associated with the penalty term (sum of squared regression coefficients), which controls the contribution of the penalty term to the cost function (sum of the squared residuals), was estimated as described in Ref. 51. These analyses were all performed in R Statistical Software version 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria). The performance of the VCR, RCR, CR, VC, RC, and CO models was compared in terms of the AUC and Brier scores computed via 10-fold CV.⁵²

The statistical significance of the difference in AUCs and Brier scores estimated via 10-fold cross validation was investigated by generating a bootstrap distribution of the difference with 300 replicates. To assess the contribution of hormone receptor information, we computed the following differences for AUC: (1) AUC of VCR–AUC of VC, (2) AUC of RCR–AUC of RC, and (3) AUC of CR–AUC of CO. We also assessed the contribution of voxel-level information to the prediction of pCR in addition to the contribution of hormone receptor information by computing the following differences for AUC: (4) AUC of VCR–AUC of RCR and (5) AUC of VCR–AUC of CR. A positive difference for each comparison implied that the first model outperformed the second model in terms of the AUC. The contributions of both the receptor information and voxel-level information were assessed using Brier scores. For each comparison, we used Brier scores *in lieu* of AUC values. Note that a negative value of the difference in Brier scores indicated that the first model outperformed the second model in terms of predictive power. P -value associated with each comparison was computed using bootstrap-estimated standard deviation (t -test).

2.3 10-Fold Cross Validation

To assess predictive power of each model corrected for overfitting, we computed AUC and Brier score via 10-fold CV. After randomly partitioning the data into 10 equally sized subgroups, each model was estimated using only nine subgroups and then validated against the remaining subgroup. After each of the 10 subgroups was used for validation, AUC and Brier score were evaluated, which completed one round of cross validation. To reduce variability, we repeated 100 rounds of 10-fold CV and averaged AUCs and Brier scores over 100 rounds, resulting in estimated AUC and Brier score for each model. We then repeated this procedure (i.e., from random partitioning to averaging AUCs and Brier scores over 100 rounds) 300 times to construct 95% confidence intervals (CI) for the estimated AUC and Brier score, with randomly resampled data with replacement. Details are described in Appendix B and Fig. 5. This approach was preferred over the leave-one-out cross validation, which tends to inflate the variance of AUC (or Brier score) estimates.⁵³

Fig. 1 (a–c) The k_{ep} and (d–f) ADC maps superimposed on the postcontrast DCE-MRI data (a and d) pre-NAT, (b and e) postone cycle of NAT, and (c and f) post all cycles of NAT for one patient achieving pCR. Multiparametric maps at different time points were all registered to a common space, which is crucial for obtaining voxel-level ADC and k_{ep} .

3 Results

3.1 Registration

In Fig. 1, representative k_{ep} (a–c) and ADC (d–f) parametric maps are superimposed on the postcontrast DCE-MRI data at baseline (a and d), and after the first (b and e) and last cycles of NAT (c and f) for a patient achieving pCR. Longitudinally acquired parametric maps at three time points were transformed into common space using the previously developed method.^{45,54} Similar data for a patient with a residual disease (i.e., a non-pCR patient) are shown in Fig. 2.

Note that for the patient achieving pCR, the mean k_{ep} within the tumor ROI decreased from 0.42 to 0.31 min^{-1} between the pre- and postfirst cycle of NAT, whereas for a patient not achieving pCR, k_{ep} increased from 0.26 to 0.32 min^{-1} between these same time points. The mean ADC increased from 1.47 to 1.75 $\text{mm}^2/\text{s} \times 10^{-3}$ between the first two time points for the same complete responder as shown in Fig. 1, while no noticeable change in mean ADC was observed for the non-pCR patient shown in Fig. 2. Both the ROI-level mean k_{ep} and ADC values were not available for the pCR patient after all cycles of NAT because there were no residual tumor voxels.

Fig. 2 (a–c) The k_{ep} and (d–f) ADC maps superimposed on the postcontrast DCE-MRI data (a and d) pre-NAT, (b and e) postone cycle of NAT, and (c and f) post all cycles of NAT for one non-pCR. Multiparametric maps at different time points were all registered to a common space.

Table 2 AUC and Brier scores (with bootstrap 95% CIs) based on 10-fold CV for the model with voxel-based information, receptor information, and clinical data (VCR model), previous model without receptor information (VC model), and model with the change in the longest dimension of the tumor, clinical data, and receptor information (RCR model), previous model without receptor information (RC model), model with clinical data and receptor information (CR model), and previous model without receptor information (CO model).

	10-fold cross validation	
	AUC	Brier score
VCR model	0.94 (0.85, 0.99)	0.11 (0.06, 0.16)
VC model	0.87 (0.71, 0.97)	0.14 (0.07, 0.20)
RCR model	0.81 (0.63, 0.95)	0.16 (0.10, 0.22)
RC model	0.71 (0.53, 0.89)	0.20 (0.13, 0.25)
CR model	0.77 (0.56, 0.93)	0.18 (0.10, 0.24)
CO model	0.67 (0.53, 0.86)	0.21 (0.15, 0.25)

However, there were residual tumor voxels showing k_{ep} (mean = 0.35 min⁻¹) and ADC (mean = 1.5×10^{-3} mm²/s) values for the non-pCR patient in Fig. 2.

3.2 Hormone Receptor Status

In Table 1, the molecular subtypes of 33 patients are summarized. Among the 33 patients, 9 (27%), 6 (18%), 8 (24%), and 10 (31%) patients are classified into luminal A, B, HER2, and basal subtype, respectively. As molecular subtypes can be determined either directly via multigeneassays or indirectly using immunochemistry, we elected to use receptor status to define the molecular subtype since this was available for all patients. We employed hormone receptor status as a covariate

in each model; importantly, the effect of each hormone receptor status on pCR status can be directly assessed by this approach.

3.3 Area Under the Curve and Brier Score via 10-Fold Cross Validation

In Table 2, the AUCs and Brier scores estimated via 10-fold CV from the six logistic ridge-regression models are reported along with the corresponding 95% CIs. The AUC from the VCR model (AUC = 0.94) outperforms all the other AUC values reported in Table 2. In particular, it outperforms the VC model, which achieved an AUC of 0.87, (i.e., about an 8% increase in the AUC compared with the VC model). The additional contribution of hormone receptor information was assessed by comparing the AUCs of the VCR, RCR, and CR models with those of the VC, RC, and CO models, achieving increases in AUC of 8%, 14%, and 15% for each comparison, respectively. It is noteworthy that the AUC of VCR also outperforms the RCR and CR models by 16% and 22%, respectively, due to using voxel-level information in addition to the hormone receptor information. Moreover, the VCR model achieves the smallest Brier score 0.11, indicating that the VCR model can predict pCR more accurately than the VC (Brier score = 0.14), RCR (Brier score = 0.16), RC (Brier score = 0.20), CR (Brier score = 0.18), and CO (Brier score = 0.21). The additional contribution of the receptor information was assessed by comparing the Brier scores of the VCR, RCR, and CR models with those of the VC, RC, and CO models achieving reductions in the Brier scores of 21%, 20%, and 14%, respectively, for each comparison. Note that the Brier score of VCR also outperforms RCR and CR models (31% and 39% improvements, respectively) due to using voxel-level information in addition to the hormone receptor information. Therefore, incorporating hormone receptor information (or molecular subtype) improves both the AUC and Brier score associated with predictive performance of the model.

The bootstrap distribution of the differences in AUCs between VCR and VC, and between RCR and RC, is shown in Fig. 3 for illustration purpose. The area of the shaded region in

Fig. 3 Differences in AUCs (a) between VCR and VC models and (b) between RCR and RC models. The area, i.e., 0.06 and 0.16, in each plot indicates how likely the AUCs based on the VCR and RCR models are smaller than the VC and RC models, respectively. Therefore, the VCR (or RCR) model would be 94% (or 84%) more likely to outperform the VC (or RC) model in terms of AUC.

Table 3 The likelihood of the first model outperforming the second for each comparison in terms of AUC and Brier scores based on 10-fold CV and the p -value associated with each comparison. Prob(AUC) and Prob(Brier score) denotes the probability that the first model (e.g., VCR) outperforms the second model (e.g., VC) in terms of AUC and Brier score, respectively. The p -values associated with t -tests for difference in AUC and Brier score between two models are computed using bootstrap-estimated standard deviation. The first three comparisons assess the contribution of the hormone receptor information to the prediction of pCR, and the last two comparisons are to investigate the contribution of voxel-level information to the prediction.

	10-fold cross validation		P -values	
	Prob(AUC)	Prob(Brier score)	AUC	Brier score
VCR versus VC	0.94	0.97	0.11	0.07
RCR versus RC	0.84	0.87	0.15	0.12
CR versus CO	0.85	0.87	0.15	0.13
VCR versus RCR	0.98	0.96	0.22	0.05
VCR versus CR	0.99	0.97	0.15	0.02

Fig. 3 (i.e., 0.06 and 0.16) indicates how the AUCs resulting from the VCR and RCR models are likely smaller than those resulting from the VC and RC models, respectively. That is, the VCR model would be 94% more likely to outperform the VC model and the RCR model would be 84% more likely to outperform the RC model in terms of AUC. Table 3 summarizes how likely the first model outperforms the second model in terms of AUCs and Brier scores for all five comparisons described in Sec. 2.2. Note that the models with additional receptor information are 84% to 94% more likely to outperform

the corresponding models without the receptor information in terms of AUC. Moreover, the models using voxel-level information (i.e., VCR) would be 98% or 99% more likely to outperform the RCR and CR models (i.e., the models do not use voxel-level information), respectively.

The bootstrap distribution of the differences in Brier scores between VCR and VC, and between RCR and RC, is shown in Fig. 4. The area of the shaded region in Fig. 4 (i.e., 0.03 to 0.13) indicates that the Brier score resulting from VCR (RCR) is larger than that from VC (RC); the VCR (RCR) model would be 97% (87%) more likely to outperform the VC (RC) model in terms of Brier score. In addition to AUCs, we have shown that the contributions of both hormone receptor information and voxel-level information are significant as displayed in the third column of Table 3. The p -values associated with comparisons for the differences in AUCs and Brier scores are summarized in the fourth and fifth columns in Table 3. None of the AUC differences is statistically significant at FDR = 0.1 except for the difference in Brier score between the VCR and CR models. Marginal significance is also observed for Brier score of the VCR versus the VC and RCR models. Although the 95% CIs reported in Table 2 and the p -values in Table 3 do not support statistically significant differences for each comparison, the values in the first two columns of Table 3 (Figs. 3 and 4) strongly suggest that the AUC and Brier scores of the VCR are likely to outperform the other models. Appendix C summarizes details on (1) the three most often chosen variables and a total number of variables chosen by redundancy analysis, (2) six estimated models with our data ($n = 33$), including penalty parameters, and (3) summary of bootstrap distribution of ridge penalty parameters.

4 Discussion

It has previously been shown³⁶ that combining spatially heterogeneous voxel-level information of quantitative MRI-derived

Fig. 4 Differences in Brier scores (a) between VCR and VC models and (b) between RCR and RC models. The area, i.e., 0.03 and 0.13, in each plot indicates how likely the Brier scores based on the VCR and RCR models are larger than the VC and RC models, respectively. Therefore, the VCR (or RCR) model would be 97% (or 87%) more likely to outperform the VC (or RC) model in terms of overfitting-corrected Brier score.

parameters (i.e., both k_{ep} and ADC) with ROI-level summary information of those parameters and clinical information (i.e., patient age and tumor grade) enhances the ability to accurately predict the response of breast tumors to NAT, compared with univariate analysis with any one parameter. Moreover, there are studies proving that tumor subtypes have had an effect on the prediction of tumor response to NAT using DCE-MRI or DW-MRI.^{37–40} Based on these developments, we have proposed a model that integrates voxel-level MRI-derived parameters with hormone receptor status (i.e., PR, ER, and HER2) in an attempt to improve the ability to predict tumor response in breast cancer patients receiving NAT. The AUC and Brier score of the proposed model, estimated via 10-fold CV (predictors: hormone receptor status and voxel-level and ROI-level k_{ep} and ADC along with clinical information) outperformed the VC (predictors: voxel-level and ROI-level k_{ep} and ADC along with clinical information), RCR (predictors: RECIST along with clinical and receptor information), RC (predictors: RECIST + clinical information), CR (predictors: clinical and receptor information), and CO (predictors: clinical information only) models.

The major limitations of the current study are the modest sample size ($n = 33$) and the lack of validation of the VCR model using an independent cohort of patients. This makes it difficult to generalize our conclusions to a population of breast cancer patients, although our initial results are very promising. Although for generalization it is imperative to have access to external sources of appropriate data to validate our findings, locating such data is quite challenging. For example, although The Cancer Imaging Archive (TCIA) is an excellent source for large imaging-based data sets, there is currently a paucity of longitudinal, quantitative, multimodal MRI data (with linked hormone receptor status) at the pre- and postone-cycle of NAT time points. Moreover, the availability of hormone receptor status is another potential limitation. In fact, to the best of our knowledge, the only appropriate data set on the TCIA is the very one analyzed in this article. The results presented in this paper speak to the importance of building publicly available data sets that include quantitative parametric maps.

Li et al.^{7,36} have shown that the combined use of two MR modalities (DCE- and DW-MRI) along with clinical and pathology data improved the prediction of response to NAT ($n = 33$, AUC ~ 0.87 with overfitting correction), although the studies that have employed only one of these modalities have also demonstrated utility in predicting tumor response to NAT.^{13,15} Park et al.⁹ have confirmed that the integration of DW-MRI and PET/CT enhanced the ability to provide accurate predictions ($n = 34$, AUC = 0.94 without overfitting correction). With mid-sized data sets ($n = 81$), Michishita and colleagues⁴⁰ have revealed that a logistic regression model that included both breast cancer subtypes and tumor morphology measured by DCE-MRI as predictors achieved an AUC of 0.83 (without overfitting correction), which outperformed the models that included only subtype or morphology alone. These results based on the small- and mid-sized data are consistent with our findings that combining all available information improves the predictive power of early tumor response to NAT. However, because adding more predictors tends to improve AUC regardless of the clinical relevance of the added information, it is imperative to correct for overfitting when evaluating statistical models.

In conclusion, we have presented a logistic ridge-regression model for predicting the response of breast tumors to NAT using

voxel- and ROI-level k_{ep} and ADC parameters, as well as hormone receptor status and clinical data. The predictive power of this model, measured with the AUC and Brier score, is superior to that of a model based only on RECIST and clinical data (i.e., the RC model) and a model based only on voxel-level parameter data and clinical data alone, without hormone receptor information (i.e., the VC model).

Appendix A: Redundancy Analysis

The goal of redundancy analysis with the set of 21 variables, $X_1, X_2, \dots, X_{21}$, is to select which are the most “nonpredictable” by the remaining variables at a given level of the coefficient of determination [in this study, we set the coefficient of determination ($= R^2$) threshold to be 0.9]. To achieve this, we first fit 21 multiple regression models. In each model, X_i (i fixed) is the dependent variable while the other 20 ($X_j, j \neq i$) are the explanatory variables. When the R^2 resulting from a model is >0.9 , then the dependent variable corresponding to the model is defined as predictable and removed from the set. If it is <0.9 , the variable is defined as nonpredictable. This procedure is repeated until no variable can be removed from the set. The remaining nonpredictable variables are defined as the most nonpredictable and are included in our prediction models, i.e., the VC and VCR models. Note that the correlation matrix among the most nonpredictable k_{ep} and ADC variables is available at Ref. 55.

Appendix B

10-Fold Cross Validation

Here, we illustrate how to compute the statistic of interest via 10-fold CV to compare among models and how to construct 95% CI on the statistic of interest. Suppose that the statistic of interest is the AUC. Then

- Step 1. Randomly partition the data into 10 (approximately) equally sized subsets
- Step 2. Use nine subsets to fit a logistic model (e.g., VCR) and compute the predicted value for each patient with the remaining subset
- Step 3. Repeat step 2 until each of 10 subsets is used for validation
- Step 4. Compute the AUC (AUC_{val}) using the observed pCRs and predicted pCRs
- Step 5. Repeat steps 1 to 4 multiple times (e.g., 100 times in this study) to reduce variability
- Step 6. Average over 100 CV rounds to establish the AUC_{cv}

Construct 95% CI on AUC_{cv}

- Step 7. Resample the data with replacement and then treat the resampled data as original data in the above-mentioned procedure
- Step 8. Repeat steps 1 to 6 described above to compute AUC_{cv} for the resampled data
- Step 9. Repeat steps 7 to 8 multiple times (e.g., 300 times in this study) to generate a bootstrap distribution of AUC_{cv}
- Step 10. Read 2.5 and 97.5 percentile values corresponding to lower and upper limit of 95% CI, respectively.

Fig. 5 Flowchart of statistical procedure to a bootstrap distribution of AUC and Brier score resulting from 10-fold cross validation.

Steps 1 to 10 can also be applied to correct for overfitting of the Brier score for each model (Fig. 5).

Appendix C: Estimated Model Parameters

The three most uncorrelated variables with the remaining variables in redundancy analysis and the total number of variables selected by redundancy analysis (Table 4).

Table 4 Three most uncorrelated variables in redundancy analysis with frequency (%) and the median number of variables chosen for VC and VCR models with 300 bootstrapped datasets.

	Max of ADC (%)	95th percentile of ADC (%)	Max of Kep (%)	Median number of variables
VC model	97.3	97.3	91.7	11 ($Q_1 = 11$, $Q_3 = 16.05$)
VCR model	98.3	97.0	91.0	11 ($Q_1 = 11$, $Q_3 = 16.06$)

Summary of estimated logistic ridge-regression models and penalty parameters using our data ($n = 33$).

Note that all the variables were centered and scaled by their standard deviations, i.e., to remove the intercept term.

1. VCR model with the ridge penalty parameter of 0.06

$$\begin{aligned}
 \text{logit}[\text{Pr}(\text{pCR} = 1)] &= -1.24 \text{ROI}_{\text{Kep}} + 2.67 \text{ROI}_{\text{ADC}} - 0.75 \text{Age} \\
 &+ 1.77 \text{Tumor} - 2.84 \text{ER} - 2.48 \text{PR} \\
 &+ 1.01 \text{HER2} - 1.32 \text{Kep}(\text{min}) \\
 &- 1.51 \text{Kep}(10\text{th}) - 1.28 \text{Kep}(95\text{th}) \\
 &- 2.11 \text{Kep}(\text{max}) - 0.77 \text{ADC}(45\text{th}) \\
 &- 1.45 \text{ADC}(55\text{th}) - 2.88 \text{ADC}(65\text{th}) \\
 &- 1.16 \text{ADC}(75\text{th}) + 0.26 \text{ADC}(80\text{th}) \\
 &- 0.36 \text{ADC}(95\text{th}) + 1.86 \text{ADC}(\text{max})
 \end{aligned}$$

2. VC model with the ridge penalty parameter of 0.07

$$\begin{aligned}
 \text{logit}[\text{Pr}(\text{pCR} = 1)] &= -1.12 \text{ROI}_{\text{Kep}} + 3.17 \text{ROI}_{\text{ADC}} - 0.61 \text{Age} \\
 &+ 2.65 \text{Tumor} - 1.34 \text{Kep}(\text{min}) \\
 &- 1.23 \text{Kep}(10\text{th}) - 1.23 \text{Kep}(95\text{th}) \\
 &- 1.15 \text{Kep}(\text{max}) - 0.74 \text{ADC}(45\text{th}) \\
 &- 1.47 \text{ADC}(55\text{th}) - 2.60 \text{ADC}(65\text{th}) \\
 &- 0.67 \text{ADC}(75\text{th}) + 0.37 \text{ADC}(80\text{th}) \\
 &- 0.24 \text{ADC}(95\text{th}) + 1.66 \text{ADC}(\text{max})
 \end{aligned}$$

3. RCR model with the ridge penalty parameter of 0.06

$$\begin{aligned}
 \text{logit}[\text{Pr}(\text{pCR} = 1)] &= -1.99 \text{Age} + 2.11 \text{Tumor} \\
 &- 2.61 \text{RECIST} - 3.11 \text{ER} \\
 &- 1.89 \text{PR} + 0.04 \text{HER2}
 \end{aligned}$$

4. RC model with the ridge penalty parameter of 0.05

$$\begin{aligned}
 \text{logit}[\text{Pr}(\text{pCR} = 1)] &= -2.09 \text{Age} + 3.13 \text{Tumor} \\
 &- 2.56 \text{RECIST}
 \end{aligned}$$

5. CR model with the ridge penalty parameter of 0.07

$$\begin{aligned}
 \text{logit}[\text{Pr}(\text{pCR} = 1)] &= 1.68 \text{Age} + 2.27 \text{Tumor} \\
 &- 2.85 \text{ER} - 1.98 \text{PR} \\
 &+ 0.13 \text{HER2}
 \end{aligned}$$

6. CO model with the ridge penalty parameter of 0.04

$$\text{logit}[\text{Pr}(\text{pCR} = 1)] = -1.88 \text{Age} + 3.84 \text{Tumor}$$

Table 5 Summary of bootstrap distribution of ridge penalty parameters.

	Q_1	Median	Q_3
VCR model	0.05	0.08	0.43
VC model	0.07	0.09	0.43
RCR model	0.06	0.07	0.08
RC model	0.04	0.04	0.05
CR model	0.06	0.07	0.08
CO model	0.04	0.05	0.06

Note that the larger the number of predictors in a model, the better the model performs in general. Therefore, we employed logistic ridge regression to avoid overfitting for the models with a large number of predictors, e.g., the VCR and VC models. The ridge penalty parameter is proportional to the number of predictors in a model, leading to fair comparisons across models with different numbers of predictors. For each bootstrapped data set, the number of predictors in the VCR and VC models changes because it depends on the result from redundancy analysis, leading to high variation (i.e., wider range) for the ridge penalty parameters corresponding to the VCR and VC models as shown in Table 5.

Disclosures

No conflicts of interest (financial or otherwise) are declared by the authors.

Acknowledgments

We would like to thank Michael Smith at Vanderbilt University Medical Center for his editorial work on this article. We thank the National Cancer Institute for funding through U01 CA142565 and R01 CA129961, and CPRIT RR160005.

References

1. S. V. Liu et al., "Neoadjuvant therapy for breast cancer," *J. Surg. Oncol.* **101**(4), 283–291 (2010).
2. "Breast cancer locally advanced (stage III)," www.nccn.org (06 April 2017).
3. T. E. Yankeelov et al., "The role of magnetic resonance imaging biomarkers in clinical trials of treatment response in cancer," *Semin. Oncol.* **38**(1), 16–25 (2011).
4. C. P. Chou et al., "Monitoring breast cancer response to neoadjuvant systemic chemotherapy using parametric contrast-enhanced MRI: a pilot study," *Acad. Radiol.* **14**(5), 561–573 (2007).
5. P. Belli et al., "Diffusion-weighted imaging in evaluating the response to neoadjuvant breast cancer treatment," *Breast J.* **17**(6), 610–619 (2011).
6. P. J. Drew et al., "Evaluation of response to neoadjuvant chemoradiotherapy for locally advanced breast cancer with dynamic contrast-enhanced MRI of the breast," *Eur. J. Surg. Oncol.* **27**(7), 617–620 (2001).
7. X. Li et al., "Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer," *Invest. Radiol.* **50**(4), 195–204 (2015).
8. L. Nilsen et al., "Diffusion-weighted magnetic resonance imaging for pretreatment prediction and monitoring of treatment response of patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy," *Acta Oncol.* **49**(3), 354–360 (2010).

9. S. H. Park et al., "Comparison of diffusion-weighted MR imaging and FDG PET/CT to predict pathological complete response to neoadjuvant chemotherapy in patients with breast cancer," *Eur. Radiol.* **22**(1), 18–25 (2012).
10. L. S. Rigter et al., "Neoadjuvant chemotherapy adaptation and serial MRI response monitoring in ER-positive HER2-negative breast cancer," *Br. J. Cancer* **109**(12), 2965–2972 (2013).
11. U. Sharma et al., "Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy," *NMR Biomed.* **22**(1), 104–113 (2009).
12. U. Tateishi et al., "Neoadjuvant chemotherapy in breast cancer: prediction of pathologic response with PET/CT and dynamic contrast-enhanced MR imaging—prospective assessment," *Radiology* **263**(1), 53–63 (2012).
13. R. Woodhams et al., "Identification of residual breast carcinoma following neoadjuvant chemotherapy: diffusion-weighted imaging—comparison with contrast-enhanced MR imaging and pathologic findings," *Radiology* **254**(2), 357–366 (2010).
14. J. Wu et al., "Intratumor partitioning and texture analysis of dynamic contrast-enhanced (DCE)-MRI identifies relevant tumor subregions to predict pathological response of breast cancer to neoadjuvant chemotherapy," *J. Magn. Reson. Imaging* **44**, 1107–1115 (2016).
15. L. M. Wu et al., "Can diffusion-weighted MR imaging and contrast-enhanced MR imaging precisely evaluate and predict pathological response to neoadjuvant chemotherapy in patients with breast cancer?," *Breast Cancer Res. Treat.* **135**(1), 17–28 (2012).
16. X. Li et al., "DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: pilot study findings," *Magn. Reson. Med.* **71**(4), 1592–1602 (2014).
17. M. L. Marinovich et al., "Early prediction of pathologic response to neoadjuvant therapy in breast cancer: systematic review of the accuracy of MRI," *Breast* **21**(5), 669–677 (2012).
18. Y. Kuroki et al., "Diffusion-weighted imaging of breast cancer with the sensitivity encoding technique: analysis of the apparent diffusion coefficient value," *Magn. Reson. Med. Sci.* **3**(2), 79–85 (2004).
19. I. K. Yirgin et al., "Diffusion weighted MR imaging of breast and correlation of prognostic factors in breast cancer," *Balkan Med. J.* **33**(3), 301–307 (2016).
20. E. A. O'Flynn et al., "Evaluating the diagnostic sensitivity of computed diffusion-weighted MR imaging in the detection of breast cancer," *J. Magn. Reson. Imaging* **44**(1), 130–137 (2016).
21. U. Sharma et al., "Role of apparent diffusion coefficient values for the differentiation of viable and necrotic areas of breast cancer and its potential utility to guide voxel positioning for MRS in the absence of dynamic contrast-enhanced MRI data," *Magn. Reson. Imaging* **30**(5), 649–655 (2012).
22. A. W. Anderson et al., "Effects of cell volume fraction changes on apparent diffusion in human cells," *Magn. Reson. Imaging* **18**(6), 689–695 (2000).
23. Y. Matsumoto et al., "In vitro experimental study of the relationship between the apparent diffusion coefficient and changes in cellularity and cell morphology," *Oncol. Rep.* **22**(3), 641–648 (2009).
24. H. Lyng, O. Haraldseth, and E. K. Rofstad, "Measurement of cell density and necrotic fraction in human melanoma xenografts by diffusion weighted magnetic resonance imaging," *Magn. Reson. Med.* **43**(6), 828–836 (2000).
25. G. Fan et al., "Usefulness of diffusion/perfusion-weighted MRI in rat gliomas: correlation with histopathology," *Acad. Radiol.* **12**(5), 640–651 (2005).
26. H. Poptani et al., "Monitoring thymidine kinase and ganciclovir-induced changes in rat malignant glioma in vivo by nuclear magnetic resonance imaging," *Cancer Gene Ther.* **5**(2), 101–109 (1998).
27. T. E. Yankeelov et al., "Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: initial results," *Magn. Reson. Imaging* **25**(1), 1–13 (2007).
28. A. R. Padhani et al., "Prediction of clinicopathologic response of breast cancer to primary chemotherapy at contrast-enhanced MR imaging: initial clinical results," *Radiology* **239**(2), 361–374 (2006).
29. M. L. Ah-See et al., "Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer," *Clin. Cancer Res.* **14**(20), 6580–6589 (2008).

30. M. D. Pickles et al., "Prognostic value of pre-treatment DCE-MRI parameters in predicting disease free and overall survival for breast cancer patients undergoing neoadjuvant chemotherapy," *Eur. J. Radiol.* **71**(3), 498–505 (2009).
 31. M. G. Heldahl et al., "Prognostic value of pretreatment dynamic contrast-enhanced MR imaging in breast cancer patients receiving neoadjuvant chemotherapy: overall survival predicted from combined time course and volume analysis," *Acta Radiol.* **51**(6), 604–612 (2010).
 32. R. Johansen et al., "Predicting survival and early clinical response to primary chemotherapy for patients with locally advanced breast cancer using DCE-MRI," *J. Magn. Reson. Imaging* **29**(6), 1300–1307 (2009).
 33. S. P. Li et al., "Use of dynamic contrast-enhanced MR imaging to predict survival in patients with primary breast cancer undergoing neoadjuvant chemotherapy," *Radiology* **260**(1), 68–78 (2011).
 34. H. J. Yu et al., "MRI measurements of tumor size and pharmacokinetic parameters as early predictors of response in breast cancer patients undergoing neoadjuvant anthracycline chemotherapy," *J. Magn. Reson. Imaging* **26**(3), 615–623 (2007).
 35. S. Mani et al., "Machine learning for predicting the response of breast cancer to neoadjuvant chemotherapy," *J. Am. Med. Inf. Assoc.* **20**(4), 688–695 (2013).
 36. X. Li et al., "Analyzing spatial heterogeneity in DCE- and DW-MRI parametric maps to optimize prediction of pathologic response to neoadjuvant chemotherapy in breast cancer," *Transl. Oncol.* **7**(1), 14–22 (2014).
 37. E. Bufi et al., "Effect of breast cancer phenotype on diagnostic performance of MRI in the prediction to response to neoadjuvant treatment," *Eur. J. Radiol.* **83**(9), 1631–1638 (2014).
 38. L. J. Grimm et al., "Can breast cancer molecular subtype help to select patients for preoperative MR imaging?," *Radiology* **274**(2), 352–358 (2015).
 39. M. Wu and J. Ma, "Association between imaging characteristics and different molecular subtypes of breast cancer," *Acad. Radiol.* **24**, 426–434 (2017).
 40. S. Michishita et al., "Prediction of pathological complete response to neoadjuvant chemotherapy by magnetic resonance imaging in breast cancer patients," *Breast* **24**(2), 159–165 (2015).
 41. C. M. Perou et al., "Molecular portraits of human breast tumours," *Nature* **406**(6797), 747–752 (2000).
 42. A. Charehbil et al., "Accuracy of MRI for treatment response assessment after taxane- and anthracycline-based neoadjuvant chemotherapy in HER2-negative breast cancer," *Eur. J. Surg. Oncol.* **40**, 1216–1221 (2014).
 43. A. C. Wolff et al., "Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update," *J. Clin. Oncol.* **31**(31), 3997–4013 (2013).
 44. M. E. Hammond et al., "American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer," *J. Clin. Oncol.* **28**(16), 2784–2795 (2010).
 45. X. Li et al., "A nonrigid registration algorithm for longitudinal breast MR images and the analysis of breast tumor response," *Magn. Reson. Imaging* **27**(9), 1258–1270 (2009).
 46. S. Rajaraman et al., "Automated registration of sequential breath-hold dynamic contrast-enhanced MR images: a comparison of three techniques," *Magn. Reson. Imaging* **29**(5), 668–682 (2011).
 47. D. A. Hamstra et al., "Evaluation of the functional diffusion map as an early biomarker of time-to-progression and overall survival in high-grade glioma," *Proc. Natl. Acad. Sci. U.S.A.* **102**(46), 16759–16764 (2005).
 48. C. R. Meyer et al., "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations," *Med. Image Anal.* **1**(3), 195–206 (1997).
 49. A. L. van den Wollenberg, "Redundancy analysis an alternative for canonical correlation analysis," *Psychometrika* **42**(2), 207–219 (1977).
 50. S. Le Cessie and J. C. Van Houwelingen, "Ridge estimators in logistic regression," *J. R. Stat. Soc. Ser. C* **41**(1), 191–201 (1992).
 51. E. Cule and M. De Iorio, "Ridge regression in prediction problems: automatic choice of the ridge parameter," *Genet. Epidemiol.* **37**(7), 704–714 (2013).
 52. B. Efron, "Estimating the error rate of a prediction rule: improvement on cross-validation," *J. Am. Stat. Assoc.* **78**(382), 316–331 (1983).
 53. P. S. L. Breiman, "Submodel selection and evaluation in regression. The x-random case," *Int. Stat. Rev.* **60**(3), 291–319 (1992).
 54. X. Li et al., "Validation of an algorithm for the nonrigid registration of longitudinal breast MR images using realistic phantoms," *Med. Phys.* **37**(6), 2541–2552 (2010).
 55. H. Kang, <http://biostat.mc.vanderbilt.edu/wiki/Main/HakmookKang> (31 October 2017).
- Hakmook Kang** received his PhD in biostatistics from Brown University. He joined the faculty at Vanderbilt University in the fall of 2011 and is an assistant professor of biostatistics. He is the core director of Biostatistics and Bioinformatics Core, Vanderbilt Kennedy Center. His research interests are in spatiotemporal modeling and multiple comparisons in biomedical imaging data analysis.
- Allison Hainline** is a PhD candidate in biostatistics at Vanderbilt University. She received her bachelor of science in statistics from Baylor University. She is interested in the application of modern statistical theory and methods to neuroimaging analysis with a focus on diffusion MRI.
- Lori R. Arlinghaus** is an imaging research scientist at Vanderbilt University Institute of Imaging Science. Her research interests focus on clinical application of quantitative magnetic resonance imaging techniques. She graduated with her BS degree in biomedical engineering from Washington University in St. Louis, 2002, and her MS and PhD degrees in biomedical engineering from Vanderbilt University, in 2005 and 2009, respectively.
- Xia Li** received her PhD in electrical engineering from Vanderbilt University in 2007. She worked as an imaging research scientist at Vanderbilt University and now is a lead scientist at GE Global Research Center. Her research interests focus on medical image processing and analysis with machine learning/deep learning.
- Vandana G. Abramson** received her MD from the University of Chicago Pritzker School of Medicine in 2000. She then completed her residency at Brigham and Women's Hospital/Harvard Medical School in 2003 and a fellowship in hematology/oncology from the University of Pennsylvania in 2006. Currently, she is an associate professor of medicine in the Department of Medicine at the Vanderbilt University Medical Center. She is the PI of numerous clinical trials and translational studies in breast cancer.
- Anuradha Bapsi Chakravarthy** completed a residency in radiation oncology (Johns Hopkins University) following completion of an internal medicine residency and a medical oncology fellowship. She is professor of radiation oncology at Vanderbilt University School of Medicine. On a national level, she serves as radiation oncology chair for the Eastern Cooperative Oncology Group as well as on the NCI-GI steering committee.
- Richard G. Abramson** received his MD degree from Harvard University and is associate professor of radiology from Vanderbilt University School of Medicine. He is a subspecialty trained diagnostic radiologist with a clinical focus in body imaging and an NIH-funded translational research portfolio centered on imaging-based response assessment for solid malignancies. He is founder and medical director of the Vanderbilt Cancer Imaging Support Laboratory and the Vanderbilt Radiology Clinical Trials Support Core.
- Kareem Fakhoury** received his BS degree in biomedical engineering from Johns Hopkins University in 2010. He is a current MD candidate at Vanderbilt University, expected to graduate in May 2018. He is interested in pursuing a career in radiation oncology with research interests spanning from biomedical imaging to integrative medicine.
- Thomas E. Yankeelov** is the W.A. "Tex" Moncrief professor of computational oncology and professor of biomedical engineering and diagnostic medicine at the University of Texas at Austin. He also serves as the director of the Center for Computational Oncology in the Institute for Computational and Engineering Sciences. The overall goal of his research is to develop tumor forecasting methods to optimize therapy by integrating advanced imaging technologies with predictive, multiscale models of tumor growth.

Biographies for the other authors are not available.