Decomposition-based framework for tumor classification and prediction of treatment response from longitudinal MRI

Rahbek, Sofie; Mahmood, Faisal; Tomaszewski, Michael R.; Hanson, Lars G. ; Madsen, Kristoffer H.

Published in:
Physics in Medicine and Biology

Link to article, DOI:
10.1088/1361-6560/acaa85

Publication date:
2023

Document Version
Peer reviewed version

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

ACCEPTED MANUSCRIPT

Decomposition-based framework for tumor classification and prediction of treatment response from longitudinal MRI

To cite this article before publication: Sofie Rahbek *et al* 2022 *Phys. Med. Biol.* in press https://doi.org/10.1088/1361-6560/acaa85

Decomposition-based framework for tumor classification and prediction of treatment response from longitudinal MRI

Sofie Rahbek[1], Faisal Mahmood[2,3], Michal R. Tomaszewski[4,5], Lars G. Hanson[1,6] and Kristoffer H. Madsen[6,7]

[1]Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark

[2]Department of Clinical Research, University of Southern Denmark, Odense, 5000, Denmark

[3]Laboratory of Radiation Physics, Department of Oncology, Odense University Hospital, Odense C, 5000, Denmark

[4]Translation Imaging Department, Merck & Co, West Point, PA, USA

[5]Cancer Physiology Department, H. Lee Moffitt Cancer Center and Research Institute, 12902 USF Magnolia Dr, Tampa, FL 33612, USA

[6]Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, 2650, Denmark

[7]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark

E-mail: khma@dtu.dk (Kristoffer H. Madsen)

Abstract. *Objective* In the field of radiation oncology, the benefit of MRI goes beyond that of providing high soft-tissue contrast images for staging and treatment planning. With the recent clinical introduction of hybrid MRI linear accelerators (MR-Linacs) it has become feasible to map physiological parameters describing diffusion, perfusion, and relaxation during the entire course of radiotherapy, for example. However, advanced data analysis tools are required for extracting qualified prognostic and predictive imaging biomarkers from longitudinal MRI data. In this study, we propose a new prediction framework tailored to exploit temporal dynamics of tissue features from repeated measurements. We demonstrate the framework using a newly developed decomposition method for tumor characterization. *Approach* Two previously published MRI datasets with multiple measurements during and after radiotherapy, were used for development and testing: T_2-weighted multi-echo images obtained for two mouse models of pancreatic cancer, and diffusion-weighted images for patients with brain metastases. Initially, the data was decomposed using the novel monotonous slope non-negative matrix factorization (msNMF) tailored for MR data. The following processing consisted of a tumor heterogeneity assessment using descriptive statistical measures, robust linear modelling to capture temporal changes of these, and finally logistic regression analysis for stratification of tumors and volumetric outcome. *Main Results* The framework was able to classify the two pancreatic tumor types with an *area under curve* (AUC) of 0.999, $P < 0.001$ and predict the tumor

volume change with a correlation coefficient of 0.513, $P = 0.034$. A classification of the human brain metastases into responders and non-responders resulted in an AUC of 0.74, $P = 0.065$. *Significance* A general data processing framework for analyses of longitudinal MRI data has been developed and applications were demonstrated by classification of tumor type and prediction of radiotherapy response. Further, as part of the assessment, the merits of msNMF for tumor tissue decomposition were demonstrated.

Keywords: prediction framework, decomposition, longitudinal MRI, radiotherapy, MR-Linac, treatment response. Submitted to: *Phys. Med. Biol.*

1. Introduction

MRI has become a crucial modality within cancer management. With the ability to deliver high soft-tissue contrast, conventional MRI techniques are part of clinical routine to enable morphological characterization of tumors. Beyond that, advanced MRI techniques that allow quantitative analysis can characterize the local tumor microenvironment. This increases accuracy in tumor grading, subtype classification and the chance to offer an optimal therapy plan [1, 2]. An example is diffusion-weighted MRI (DWI), which is sensitive to the local cellular density, potentially important for identification of sub-regional tumor load and for early evaluation of a tumor's response to therapy [3]. This is due to the response to ionizing radiation at a cellular level occurring on a much shorter timescale (hours to days) than the volumetric response [4]. With an early response assessment there is time to adjust the therapy plan to increase treatment efficacy or reduce unnecessary radiation [5, 6, 7, 8].

The recent introduction of hybrid MRI linear accelerator systems (MR-Linacs) [9, 10] facilitates daily MRI-based adaption of the dose plan and offers opportunities for longitudinal MRI studies of therapy response within the normal clinical workflow. However, there is a need for more custom analysis strategies for the growing pool of longitudinal MRI data. Currently, only few previous human studies include more than two measurements during treatment [8], and even fewer datasets include daily measurements [11, 12]. An analysis pipeline customized for longitudinal datasets requires a method for investigating temporal tumor dynamics, but also a procedure for extracting quantitative information from multi-contrast MRI data (consisting of three spatial dimensions and one b-value dimension for diffusion-weighting, for example). Most previous studies have utilized a model-based parametrization of the data for quantification and e.g. showed correlations between the DWI-derived apparent diffusion coefficient (ADC) and treatment outcome, e.g. [13, 14], or between dynamic contrast-enhanced (DCE) MRI parameters and tumor subtype and prognosis, e.g. [15, 16]. Model-based analyses, although useful, carry the risk of introducing biased or misleading estimates if model assumptions are wrong, e.g due to partial-volume effects. Novikov *et al* [17] outlines pitfalls of modelling tissue microstructure, and Satta *et al* [18] highlights contradicting correlations found across studies that analyzed DWI- and DCE-derived parameters, demonstrating the problem and hence a need for alternatives such as data-driven analysis.

Among the few existing studies that included longitudinal measurements in their correlation analysis, most of them investigated the temporal changes only by pairwise comparisons [19, 20, 21, 22, 23, 24]. This results in extensive multiple comparisons problems for studies involving more than 2-3 measured time points. Additionally, data trends unfolding over several time-points will be difficult to identify, though they may be important in a therapy outcome prediction, for example. Advanced models such as neural networks has been used to discover patterns across time-resolved measurements [25, 26], but these analyses were voxel-based and thus depended on an exact

image registration between measurements. This may be problematic if morphological changes occur during the course of disease.

Therefore, we propose an analysis framework which avoids both model-based parametrization for quantitative mapping and the need for voxel correspondence over time [7]. The framework provides all processing steps necessary to analyze and exploit the dynamics across longitudinal measurements for a final prediction. As an alternative to model-based analysis, the framework utilizes a data-driven blind-source separation technique to decompose the MRI data. Though any decomposition technique in principle can be used, we suggest and explore the recently published monotonous slope non-negative matrix factorization (msNMF) [27]. The msNMF is specifically tailored for decomposition of MR signals and has proven useful for extracting realistic components informative of the underlying tissue structure.

The aim of this paper is to present a full prediction framework in which longitudinal MRI data is analyzed in order to extract information with potential prognostic or predictive value in cancer management, and to explore the msNMF as the initial decomposition method for tumor tissue characterization. We use two previously analysed and published data sets, T_2-weighted multi-echo data for pancreatic tumor type classification in mice, and DWI of human brain metastases for outcome prediction.

2. Methods

2.1. The prediction framework

An overview of the prediction framework is summarized in Figure 1a, and described below.

2.1.1. Data-driven decomposition As a first step, the data matrix was separated into latent signal components using msNMF, which is an extension of standard non-negative matrix factorization (NMF). It constrains the signal components and their slopes (first derivates) to be monotonous, thus enforcing near-exponential signal decays consistent with prior expectations for the relevant data types (e.g. DWI or relaxometry). A detailed description of the msNMF was published by Rahbek *et al* [27]. For brevity, only the resulting optimization problem is summarized here:

$$\mathcal{C} = \|\mathbf{X} - \mathbf{WH}\|_F^2 \tag{1}$$

The norm of the residual, $\mathcal{C}$, is minimized, where $\|\cdot\|_F^2$ denotes the Frobenius norm. $\mathbf{X} \in \mathbb{R}_+^{m\times n}$ is the data matrix for m measurement points in a scan session (e.g. echo times or b-values) and n sources (e.g. voxels). $\mathbf{WH}$ is the low-rank representation of the data with $\mathbf{W} \in \mathbb{R}_+^{m\times k}$ being the k fundamental signal profiles and $\mathbf{H} \in \mathbb{R}_+^{k\times n}$ the associated spatial distributions ("mixture map"). $\mathcal{C}$ is minimized using an alternating non-negative least squares (ANLS) algorithm, i.e. by optimizing one factor while keeping

the other fixed. **H** is optimized under a non-negativity constraint, and **W** is optimized under both a non-negativity constraint and the additional monotonicity constraints. Strategies demonstrated by Rahbek *et al* [27] were followed regarding implementation including initialization, stopping criteria and rank determination.

2.1.2. Tumor heterogeneity assessment The decomposition output, k mixture maps describing the spatial distributions of different signal features across the volume, was analyzed for assessment of tumor features using a set of first-order radiomics. Five percentiles, (0.1, 0.25, 0.5, 0.75 and 0.9), were used to characterize the intensity distribution across the tumor volume, defined by a region-of-interest (ROI) for each component. ROI histogram values have previously been demonstrated as relevant descriptors of tumor structures [5, 7, 28, 29]. Additionally, they are robust to outliers and imperfect image registrations as opposed to voxel-based values and the ROI mean. The quartiles (0.25, 0.5, and 0.75) were supplemented with bins for extreme values, (0.1 and 0.9), to capture features of non-normal distributions while disregarding outliers, if such are present. The ROI delineation is explained in coming subsections dedicated to the individual data sets.

2.1.3. Temporal dynamics analysis The five component percentiles were calculated for all subjects and scan time-points. To capture a trend across time-resolved measurements, a linear fit was made for each percentile as a function of days after first treatment. A Huber loss function was used in the regression for the fit to reduce influence of potential outliers [30]. The baseline scan (obtained prior to treatment) was left out from the fit. Instead, the difference between the values at baseline and the first scan after treatment onset was calculated to include the immediate treatment response explicitly in the prediction model. Very early ADC decrease following RT onset has, for example, previously been reported [31] in responding brain metastases. The decrease may indicate cell swelling, an early state of cell death and therefore a potential predictor of response [11, 32]. The final prediction features then constituted the slope and intercept of the linear fit and this "difference to baseline" for each of the five percentiles generated for each of the components. An example is presented in Figure 2 using one of the observations in the DWI dataset. Lastly, each feature was normalized to the range [-1 1], such that all features were weighted equally in the prediction analysis.

2.1.4. Feature selection and prediction modelling The above processing resulted in a large number of features compared to the number of samples. To avoid overfitting and numerical instability [33, 34], the prediction analysis was thus carried out using an integrated feature selection process. Logistic regression was used as statistical model for prediction of classes and multiple (linear) regression for prediction of continuous outcomes. A nested cross-validation (CV) procedure was implemented for model training to avoid information leakage into the test data. In the outer CV loop, stratified K-fold cross-validation was used to divide data into a number of randomly stratified

groups equal to the number of samples in the least frequent class for the classification analysis. For the regression analysis, leave-one-out cross-validation was used, resulting in one sample per group. As the total number of samples were low, this setting was used to ensure that the maximum number of samples were available for training. The inner CV loop then served to determine which features were most relevant for prediction, and the input data was thus again divided into test and training data (Figure 1b). To this end, we relied on a greedy approach where the features were ranked according to test statistics from a univariate Student's t-test with pooled variance within the remaining groups. The performance estimated in the inner CV loop was used to determine the optimal threshold, and the model was retrained on the entire training set with the optimal threshold. As we did not use regularization in the current analysis, the maximum number of features was set to 10 to avoid overfitting and rank deficiency. The final generalization performance was evaluated in terms of the area under the receiver operating characteristic curve (AUC) or the Pearson correlation coefficient (ρ). The significance of the performance metrics was evaluated by a permutation test where an empirical null distribution was formed by repeating the entire analysis 5000 times with randomly permuted labels, which is a commonly used method for validating prediction performance [35]. A threshold of $\alpha = 0.05$ defined statistical significance.

[Figure 1 about here.]

[Figure 2 about here.]

2.2. T_2-weighted multi-echo imaging of mice

This dataset was previously published by Tomaszewski *et al* [36] and consists of MR imaging of mice grafted subcutaneously with either slow growing human pancreatic adenocarcinoma BXPC3 (N=8) or more aggressive mouse pancreatic adenocarcinoma Panc02 tumors (N=10). A multi-slice multi-echo sequence with 32 echoes and 7 ms echo spacing was used. The sequence was part of a full MRI protocol applied two hours prior to irradiation and every 3 days after, using a 7 T animal scanner (Bruker horizontal bore, running Paravision 6.0.1). The final imaging was performed at day 9 and day 12, respectively, for the Panc02 and BXPC3 group. The radiotherapy consisted of a single dose of 10 Gy irradiation, matching the dose per fraction of a clinical stereotactic body radiation protocol. The events of imaging and radiation is schematically illustrated in Figure 3a.

For each mouse and scan session, the tumor region, i.e. the region-of-interest (ROI), was delineated on a high-resolution T_2-weighted anatomical scan included in the full protocol. The grafted tumors were clearly visible with distinct borders, and the ROIs were therefore delineated on non-aligned images. The ROI voxels for all mice and all scan days were pooled in a single [$N_{echoes} \times N_{voxels}$] matrix for the msNMF analysis. After estimation of the signal components ($\mathbf{W}$) using this pooled data matrix, a projection of the individual data sets onto $\mathbf{W}$ resulted in mixture maps ($\mathbf{H}$) per subject and scan

session, making it possible to follow the temporal changes. The data was normalized to start at intensity 1 to mitigate effects of irrelevant signal behavior across the brain, e.g. stemming from receive coil inhomogeneities, which otherwise would influence the decomposition.

In the source paper [36], Tomaszewski *et al* calculated model-based T_2-values for each voxel in the tumor ROI and showed (statistical) differences between the two tumor types. Additionally, the paper revealed correlations between the T_2 inter-quantile range and the therapy response defined as the tumor volume change measured between the two last MRI scans. In this study, the data was thus ideal to test the proposed framework for both a classification of the two tumor types and a prediction of the therapy response. For the therapy response prediction, the BXPC3 and Panc02 groups were pooled together due to their small sample sizes.

2.3. DWI of brain metastases

This clinical dataset was previously analyzed by Mahmood *et al* [11, 31, 37] and consists of MRI brain scans of cancer patients undergoing palliative radiotherapy. The data was recorded using a 1 T MR system (Panorama, Philips Healthcare, The Netherlands). In addition to T_1- and T_2-weighted imaging, the MRI-protocol included an echo-planar imaging (EPI) DWI sequence with eight b-values of [0, 50, 100, 150, 400, 500, 600, 800] s/mm^2, three orthogonal diffusion-weighting directions, an effective diffusion time $\sim$49 ms, and a prior spectral inversion module (SPIR) for fat-suppression. The radiotherapy consisted of 10 fractions of 3 Gy whole-brain irradiation delivered over two weeks. The MRI scan sessions took place prior to, during, and after the treatment period (Figure 3b).

For each patient, one to three brain metastases were delineated by a radiologist using b=800 s/mm^2 images, guided by the T_2-weighted data. All scans were aligned with the baseline scan using rigid registration (SPM12 Statistical Parametric Mapping software, version 7487 [38]) such that the baseline tumor delineation could define the ROI for all scans. However, ROIs were expanded using a morphological dilation of 3 voxels (5.4 mm). This was both to include sub-clinical disease, edema etc. and to increase robustness to imperfect image registrations or inaccurate tumor delineation. ROI voxels were pooled for all metastases, i.e. from all subjects and scan days, resulting in a [N_{voxels} $\times$ $N_{bvalues}$] data matrix for the msNMF, although input signals largely affected by noise were discarded. These voxels were defined as having maximum signal for a non-zero b-value. The data was normalized to start at an intensity of 1.

The full framework was tested for prediction of the local control evaluated at a follow-up scan obtained 2-3 months after last irradiation. The tumor volume change from baseline to follow-up, based on a radiologist's delineation on a high-resolution T_1-weighted scan, was used to divide the metastases into responders (more than 30 % volume shrinkage) and non-responders (less than 30 % volume shrinkage). Outcome categorization was validated by the tumor ADC at follow-up to minimize the risk of

misinterpreting pseudo-progression. A total of 31 metastases (23 responders and 8 non-responders) distributed in 15 patients were analysed.

[Figure 3 about here.]

3. Results

3.1. T_2-weighted multi-echo imaging of mice

3.1.1. Decomposition with msNMF An inspection of the decomposition of ranks two to four resulted in a decomposition into three components, corresponding to 99.4% explained data variance. Figure 4a shows the resulting signal components, all being relatively smooth, realistic signal decays. The associated mixture maps (**H**) are presented in Figure 4b for a representative mouse from each group (BXPC3 and Panc02) with the tumor ROIs shown in red. Even though both tumor types predominantly contain the intermediate decaying signal component (yellow), it is clear from the images that the two tumors differ in tissue heterogeneity. Looking at the mixture maps for the long-lived signal component (blue), for example, only the BXPC3 tumor contains a sub-area with relatively high intensity.

[Figure 4 about here.]

3.1.2. Classification of the tumor type Due to the randomization included in the K-fold CV procedure, running the prediction framework 50 times resulted in a mean AUC score of 0.999, i.e. an almost perfect classification of the two tumor types. Figure 5 shows this score relative to the null distribution of 5000 runs with permuted class labels. The corresponding p-value was equal to 0.0004, which confirms that the classifier has learned a significant class structure and could distinguish the two tumor types.

Except for a single feature related to the green signal component, the blue long-lived signal component was the only relevant predictor in a classification of the tumor type (Figure 5). Four percentiles are represented for this component distribution, and both changes in intercepts and slopes explain a significant part of the variance. The latter indicates that the signal characteristics of the two tumor types differ with respect to both starting point and changes following radiation.

[Figure 5 about here.]

3.1.3. Prediction of volume change The two groups were pooled together to a total of 18 samples for the prediction of the tumor volume change, i.e. the growth between day 6 and 9 for Panc02 tumors and between day 9 and 12 for BXPC3 tumors. A correlation between the true and predicted values resulted in a correlation score $\rho = 0.513$ (Figure 6b), which compared to the null-distribution corresponded to a p-value of 0.034 (Figure 6a) reaching statistical significance. Again, features of the blue long-lived signal component were clearly of most importance to the prediction, especially the 90th percentile intercept and the 25th percentile slope (Figure 6c).

[Figure 6 about here.]

3.2. DWI of brain metastases

3.2.1. Decomposition with msNMF Considering the mix of tumor and sub-clinical tissues included in the ROIs, it was expected that at least three signal components were distinguishable. The data was thus inspected and found to contain three to five components, and as Figure 7 reveals, a rank of four was chosen for the final decomposition. The four components were able to explain 99.1 % of the data variation, and their behavior differed from each other both as a function of b-value (Figure 7b) and spatially (Figure 7c). The close-ups in Figure 7c show that for this example, the main tumor region (central part of the magenta ROI) primarily contains the two intermediate decaying signal components (yellow and green) while a different signal composition is seen for the surrounding abnormal tissues. The very rapidly decaying signal component (red) shows no contrast in the presented slice except for a few small spots of high intensity which could well be blood vessels. The intensity may be affected by flow but the attenuation of blood signal with b-value will be strong in any case.

3.2.2. Prediction of therapy response For the classification of responders versus non-responders 50 prediction models were generated due to the randomization included in the K-fold CV procedure. This resulted in a mean AUC score of 0.74, which compared to the null distribution corresponded to a p-value of 0.065 (Figure 7). The relation between tumor features and therapy outcome was thus not strong enough for the prediction to be significant at an $\alpha = 0.05$ level. The most informative feature was the start median (P50 intercept) of the b=0 s/mm^2 data (Figure 6).

[Figure 7 about here.]

[Figure 8 about here.]

4. Discussion

The main purpose of this study was to present and demonstrate a full analysis framework utilizing longitudinal MRI measurements for prognosis and prediction of cancer treatment outcome. Our work was motivated by the recent release of MR-Linacs, which is expected to increase the amount of longitudinal MRI cancer studies in the coming years. Additionally, the potential problems and pitfalls connected with model-based analysis pointed out by, e.g., Novikov *et al* [17] and Satta *et al* [18] motivated a framework that avoids model-based parametrization. To show its flexibility and range of applications, the framework was demonstrated using two datasets that varied both regarding MR contrast, number of contrast measurements, and number of time-resolved scans. Though results were statistically significant only for the T_2-weighted multi-echo dataset, the framework could handle both types of data.

4.1. T_2-weighted multi-echo imaging of mice

The preclinical data published by Tomaszewski *et al* [36] was ideal to use for framework testing. Firstly, it included a set of well-resolved signals (short echo spacing) measured multiple times during the treatment period. Secondly, the analyses of Tomaszewski *et al* revealed interesting findings regarding the predictive value of the T_2 inter-quantile range (reflecting tumor heterogeneity), which were supported by histology. Finally, the two pancreatic tumor models allowed to test the framework's ability to do a simpler tumor classification before testing for prediction of outcome.

The three signal components detected by the msNMF analysis clearly presents different relaxation features of the tumor tissue. In the raw data, decay curves contained signal transients from RF inhomogenety in the beginning of the echo train appearing as small oscillations, but the components are unaffected due to the msNMF constraints (Supplementary Figure S1). As discussed in Rahbek et al [27], the decomposition is tailored to give physically meaningful components. In some cases, however, the data may not be sufficiently informative to distinguish the true (underlying) signal components and different solutions may explain the data almost equally well. Nevertheless, the resulting decompositions may well carry predictive value.

The prediction analysis was able to distinguish between the two tumor types with very high accuracy, and the successful stratification largely depended on tissue information involving the long-lived signal component. This agrees with results of the source paper [36] revealing that the BXPC3 tumor has a broader distribution of T_2-values including a longer T_2 relative to the Panc02 tumor, though direct comparisons are difficult since standard T_2 estimates are compromised by partial-volume effects. Satisfactorily, histology data confirmed a lower cellular density and more heterogeneous tissue for the BXPC3 tumor, consistent with the detected prominence of a long-lived fluid signal [39]. Although the convincing histological difference between the two tumor types allowed a robust classification to be expected, the results confirm that relevant tissue information is preserved throughout the comprehensive processing and used in the classification.

The analysis was also able to predict the tumor volume change with a significance of p=0.034 compared to the null distribution, a satisfactory prediction considering the low number of samples for both training and testing as well as the heterogeneity introduced by the use of two distinct tumor models. Again, the long-lived signal component was crucial, both the estimated starting point and the rate of change. One explanation could be that these values relate to the level of cell death and formation of necrosis (supported by the histological analysis), which naturally alters the local T2-weighted relaxation. Specifically, T_2-values decrease with increasing cell density [39]. While necrosis is an expected response to radiation, the level of necrosis before treatment can conversely also affect the response [40, 41]. Tomaszewski *et al* found necrosis to be

reflected in a heterogeneity measure (the T_2 interquartile range) [36]. Our results point at the abundance of the long-lived signal component as a possible indicator of necrosis. We also tested the predictive value of the estimated T_2-values by running the framework using the T_2-maps instead of the mixture maps from the decomposition. This led to a significant prediction with a correlation of 0.62 and p=0.0065, where the most important feature was the intercept of the 90th percentile of T_2s. Again, this points to a relevant correlation between the long-lived T_2-weighted decay and the treatment response.

Due to limited data, it was necessary to pool the two tumor groups for the outcome prediction. However, underlying biological differences may make it difficult to find a general correlation between the data and the tumor volume response. For example, cell swelling, the extent of which depends on the tumor type, may occur during the initial phase of the therapy-induced apoptosis [7, 42]. If this or other radiobiological traits differ between Panc02 and BXPC3 tumors, it may explain why features of the immediate radiation response were not relevant to the prediction model. Tomaszewski et al supports this, as they reported that slowdown of tumor growth rate was detected already as early as day 3 for the Panc02 tumors but not until day 6 for the BXPC3 tumors. Nevertheless, the use of two groups in one model makes the solution more generalizable, even if the model performance is not as good. It is highly encouraging that a significant prediction of the tumor volume change was obtained despite of variations in tumor type.

The analysis of the mice relaxometry data has confirmed the potential of our framework and revealed that the signal component with a slow decay (long T_2) was important in terms of successfully classifying tumor type and predicting tumor volume change. We cannot expect it to be a general finding, but the meaningful biological link to necrosis and high heterogeneity is interesting. Larger datasets are necessary to robustly identify signal trends and relate these to tissue features.

4.2. DWI of brain metastases

The DWI dataset previously presented by Mahmood *et al* [11] is rather unique as it contains many b-values and consists of scans acquired at each of the ten fractions during the radiotherapy course. These properties made the dataset suited for testing of the proposed framework, although the sample size is too small to expect reliable predictions. Additionally, the work by Mahmood *et al* [31, 37] showed differences between responders and non-responders with respect to the relative ADC change during therapy and thus indicated that DWI signals carry information useful for early stratification of treatment response, in consistency with other studies [8, 43, 20].

The prediction was close to the threshold of significance (AUC=0.74, p=0.065) indicating a possible relationship between the data and the response labels. The heterogeneity of the data could be one reason for the non-significance. The metastasis

originated from six different primary tumor types (listed in Mahmood *et al* [37]), and excessive variation can make it difficult to establish a general relationship between the signal signature and the treatment outcome with limited data. The driving feature in the prediction originated from the b=0 s/mm^2 data. A prediction test that included only these T_2-weighted data led to a very similar result (AUC=0.74, p=0.052), indicating that local T_2-relaxation was relevant, while diffusion-weighted signal did not provide additional information for the response stratification in this case. It should be emphasized that demonstrating group differences of the ADC as in [31, 37] does not necessarily imply that a general pattern from input data to output label can be learned. Inter- and intra-tumor heterogeneity may dominate the variation in the diffusion-weighted data. We tested the predictive value of the ADC-values directly by running the framework using ADC-maps instead of mixture maps and b=0 s/mm^2 data. As expected, this resulted in a poor prediction with AUC=0.55 and p=0.36.

There is a notable limitation connected to the use of numerical bounds to separate and define non-responders from responders as was the case for this data. The RECIST guidelines [44] were followed except for using tumor volume instead of tumor diameter (as in [31, 37]). The dichotomization becomes particularly problematic here because several observations out of the small cohort (N=31) had a relative volume change close to the threshold. We also attempted prediction of the actual volume change using a regression model instead, but this resulted in a poor prediction outcome (analysis not shown). Despite the highlighted data challenges, it was demonstrated that the framework is well-suited for longitudinal study designs and relevant data types, which was a main objective.

A challenge not specific to this data, but more general for studies relying on tumor features, involves the definition of the tumor ROI, for example which images (DWI, T_2-weighted,...) to use for delineation and which target volume (gross tumor volume, clinical target volume, viable tumor volume, etc.) to choose [45, 46, 47]. Mahmood et al showed how the ROI delineation strategy in conjunction with the signal fitting method affected the estimated ADC changes [11]. This is one of the reasons that our framework avoids model-based parameter estimation and is designed to be less sensitive to ROI variability. The unsupervised component analysis is not confounded by intra-voxel heterogeinity, and the ROI dilation and histogram-based assessment of tumor features reduce the significance of exact delineations. Nevertheless, there is still reason to be critical towards the choice of target and ROI dilation. We utilized the delineations from [31], i.e. high-intensity regions at b=800 s/mm^2 corresponding to "viable" tumor dilated by approximately 5 mm, and obtained reasonable results of the framework. These choices deserve investigation in studies with more data, and optimally an automatic segmentation process can eventually replace the manual delineation.

For the preclinical data example, the mice tumors, which are generated from subcutaneous cell injections into the hind leg, resulted in clearly delimited large tumor regions

throughout the course of monitoring, why the ROI strategy was not an issue there.

An important aim of the proposed framework is the utilization of information from all scans simultaneously. Previous longitudinal studies assumed independence between the data points and treated them individually, although being related to the same patient. We argue it is better data handling to include the measurements' dependency, especially as there may be predictive value of the early changes for the later therapy outcome.

Also, with the improved availability of MR-Linacs there is less reason to settle for a single scan time point. A robust linear fit has been proposed for this processing step because it is less sensitive to small variations (noise) in the data than a fit of higher order, which also introduces ambiguity of fitted coefficients when linear fitting is sufficient. Linear fitting is also more robust towards missing measurements and which time-points that are part of the longitudinal scan scheme. It is, however, possible to use more flexible modelling in the temporal dynamics analysis, if this can be trusted to generalize across observations without overfitting.

The proposed framework for analyzing longitudinal MRI datasets for tumor disease management was presented and demonstrated in a structured manner (Figure 1), highlighting all processing steps necessary to go from high-dimensional MR data to a final prediction including an unbiased performance evaluation. The cornerstones of this unique framework are the data-driven signal decomposition that avoids model-based parametrization, and the histogram-based tumor heterogeneity assessment, where tracking of individual voxels is unnecessary. Simultaneously, the framework is flexible and allows the user to replace some of the suggested methods, including the choice of data-driven decomposition, ROI definition, and temporal dynamics analysis. The decomposition strategy and its associated assumptions regarding the data structure must be appropriate for the particular data. The unique constraints of the msNMF made it an ideal choice for the given data examples. The fact that the framework functioned as anticipated and even delivered convincing results for the mice data reinforces the choices made here, and proves msNMF to be a potentially relevant decomposition technique for identifying tumor-specific signal features. The two demonstration examples indicated a value of the sources of tissue contrast to stratify tumors and predict therapy response. These results show the potential of the method for detection of biomarkers sensitive to e.g. treatment-induced changes.

For research studies, it is in principle advantageous to use independently suggested analysis frameworks to avoid confirmation biases, but the analysis should also match the particular data at hand. The diverse data of the example studies demonstrated that the framework is sufficiently versatile, and thus a relevant candidate in many cases. Overall, there are no limitations to the type of input features in the prediction, which can also be a mix of e.g. imaging and biometric features.

5. Conclusion

Our study presented a new prediction framework developed for analysis of MRI data from repeated measurements and showed its potential as a tool to identify possible MRI biomarkers, specifically within the field of cancer treatment. In particular, the framework is potentially suited for the increasing volume of longitudinal MRI data already coming from clinical trials with the recent introduction of MR-Linacs in radiation therapy. Furthermore, the merits of the newly developed msNMF for tumor tissue signal decomposition was demonstrated as part of the assessment. Applications were exemplified by classification of tumor tissue and prediction of therapy outcome using T_2-weighted multi-echo data and DWI data, respectively, demonstrating the wide applicability of the framework.

Acknowledgements

The authors would like to thank Dr. Robert J. Gillies for discussions. The study was funded by the Danish Cancer Society (grant R167-A10637-17-S2).

6. References

[1] Hatef Mehrabian, Jay Detsky, Hany Soliman, Arjun Sahgal, and Greg J. Stanisz. Advanced magnetic resonance imaging techniques in management of brain metastases. *Frontiers in Oncology*, 9(440):1–16, 2019.

[2] Frederic G. Dhermain, Peter Hau, Heinrich Lanfermann, Andreas H. Jacobs, and Martin J. van den Bent. Advanced MRI and PET imaging for assessment of treatment response in patients with gliomas. *The Lancet Neurology*, 9(9):906–920, 2010.

[3] Evelyn Dappa, Tania Elger, Annette Hasenburg, Christoph Düber, Marco J. Battista, and Andreas M. Hötker. The value of advanced MRI techniques in the assessment of cervical cancer: a review. *Insights into Imaging*, 8(5):471–481, 2017.

[4] Stephen J. McMahon and Kevin M. Prise. Mechanistic modelling of radiation responses. *Cancers*, 11(2):1–23, 2019.

[5] Manijeh Beigi, Anahita Fathi Kazerooni, Mojtaba Safari, Marzieh Alamolhoda, Mohsen Shojaee Moghdam, Shiva Moghadam, Hamidreza SalighehRad, and Ahmad Ameri. Heterogeneity analysis of diffusion-weighted MRI for prediction and assessment of microstructural changes early after one cycle of induction chemotherapy in nasopharyngeal cancer patients. *Radiologia Medica*, 123(1):36–43, 2018.

[6] DW Zhao, WJ Fan, LL Meng, YR Luo, J Wei, K Liu, G Liu, JF Li, X Zang, M Li, XX Zhang, and L Ma. Comparison of the pre-treatment functional MRI metrics' efficacy in predicting Locoregionally advanced nasopharyngeal carcinoma response to induction chemotherapy. *Cancer Imaging*, 21(59):1–12, 2021.

[7] Anwar R. Padhani, Guoying Liu, Dow Mu-Koh, Thomas L. Chenevert, Harriet C. Thoeny, Taro Takahara, Andrew Dzik-Jurasz, Brian D. Ross, Marc Van Cauteren, David Collins, Dima A. Hammoud, Gordon J.S. Rustin, Bachir Taouli, and Peter L. Choyke. Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. *Neoplasia*, 11(2):102–125, 2009.

[8] Petra J. van Houdt, Yingli Yang, and Uulke A. van der Heide. Quantitative Magnetic Resonance Imaging for Biological Image-Guided Adaptive Radiotherapy. *Frontiers in Oncology*, 10(615643):1–9, 2021.

[9] B. W. Raaymakers, I. M. Jürgenliemk-Schulz, G. H. Bol, M. Glitzner, A. N.T.J. Kotte, B. van Asselen, J. C.J. de Boer, J. J. Bluemink, S. L. Hackett, M. A. Moerland, S. J. Woodings, J. W.H. Wolthaus, H. M. van Zijp, M. E.P. Philippens, R. Tijssen, J. G.M. Kok, E. N. de Groot-van Breugel, I. Kiekebosch, L. T.C. Meijers, C. N. Nomden, G. G. Sikkes, P. A.H. Doornaert, W. S.C. Eppinga, N. Kasperts, L. G.W. Kerkmeijer, J. H.A. Tersteeg, K. J. Brown, B. Pais, P. Woodhead, and J. J.W. Lagendijk. First patients treated with a 1.5 T MRI-Linac: Clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment. *Physics in Medicine and Biology*, 62(23):L41–L50, 2017.

[10] S Mutic, D Low, T Chmielewski, G Fought, M Hernandez, I Kawrakow, A Sharma, S Shvartsman, and J Dempsey. TU-H-BRA-08: The Design and Characteristics of a Novel Compact Linac-Based MRI Guided Radiation Therapy (MR-IGRT) System. *Medical Physics*, 43(6):3770–3770, 2016.

[11] Faisal Mahmood, Helle H. Johannesen, Poul Geertsen, Giske F. Opheim, and Rasmus H. Hansen. The effect of region of interest strategies on apparent diffusion coefficient assessment in patients treated with palliative radiation therapy to brain metastases. *Acta Oncologica*, 54(9):1529–1534, 2015.

[12] T. Bostel, C. Dreher, D. Wollschläger, A. Mayer, F. König, S. Bickelhaupt, H. P. Schlemmer, P. E. Huber, F. Sterzing, P. Bäumer, J. Debus, and N. H. Nicolay. Exploring MR regression patterns in rectal cancer during neoadjuvant radiochemotherapy with daily T2- And diffusion-weighted MRI. *Radiation Oncology*, 15(1):1–12, 2020.

[13] Bo Zhao, Kun Cao, Xiao Ting Li, Hai Tao Zhu, and Ying Shi Sun. Whole lesion histogram analysis

of apparent diffusion coefficients on MRI predicts disease-free survival in locally advanced squamous cell cervical cancer after radical chemo-radiotherapy. *BMC cancer*, 19(1):1115, 2019.

[14] Masamitsu Hatakenaka, Katsumasa Nakamura, Hidetake Yabuuchi, Yoshiyuki Shioyama, Yoshio Matsuo, Kayoko Ohnishi, Shunya Sunami, Takeshi Kamitani, Taro Setoguchi, Takashi Yoshiura, Torahiko Nakashima, Kei Nishikawa, and Hiroshi Honda. Pretreatment apparent diffusion coefficient of the primary lesion correlates with local failure in head-and-neck cancer treated with chemoradiotherapy or radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 81(2):339–345, 2011.

[15] Ingfrid S. Haldorsen, Renate Grüner, Jenny A. Husby, Inger J. Magnussen, Henrica M.J. Werner, Øyvind O. Salvesen, Line Bjørge, Ingunn Stefansson, Lars A. Akslen, Jone Trovik, Torfinn Taxt, and Helga B. Salvesen. Dynamic contrast-enhanced MRI in endometrial carcinoma identifies patients at increased risk of recurrence. *European Radiology*, 23(10):2916–2925, 2013.

[16] Juliette A. Loncaster, Bernadette M. Carrington, Johnathan R. Sykes, Andrew P. Jones, Susan M. Todd, Rachel Cooper, David L. Buckley, Susan E. Davidson, John P. Logue, Robin D. Hunter, and Catharine M.L. West. Prediction of radiotherapy outcome using dynamic contrast enhanced MRI of carcinoma of the cervix. *International Journal of Radiation Oncology Biology Physics*, 54(3):759–767, 2002.

[17] Dmitry S. Novikov, Valerij G. Kiselev, and Sune N. Jespersen. On modeling. *Magnetic Resonance in Medicine*, 79(6):3172–3193, 2018.

[18] Serena Satta, Miriam Dolciami, Veronica Celli, Francesca Di Stadio, Giorgia Perniola, Innocenza Palaia, Angelina Pernazza, Carlo Della Rocca, Stefania Rizzo, Carlo Catalano, Silvia Capuani, and Lucia Manganaro. Quantitative diffusion and perfusion MRI in the evaluation of endometrial cancer: validation with histopathological parameters. *The British Journal of Radiology*, 94(1125):1–12, 2021.

[19] Ramesh Paudyal, Jung Hun Oh, Nadeem Riaz, Praveen Venigalla, Jingao Li, Vaios Hatzoglou, Jonathan Leeman, David Aramburu Nunez, Yonggang Lu, Joseph O. Deasy, Nancy Lee, and Amita Shukla-Dave. Intravoxel incoherent motion diffusion-weighted MRI during chemoradiation therapy to characterize and monitor treatment response in human papillomavirus head and neck squamous cell carcinoma. *Journal of Magnetic Resonance Imaging*, 45(4):1013–1023, 2017.

[20] Warren D. Foltz, Andy Wu, Peter Chung, Charles Catton, Andrew Bayley, Michael Milosevic, Robert Bristow, Padraig Warde, Anna Simeonov, David A. Jaffray, Masoom A. Haider, and Cynthia Ménard. Changes in apparent diffusion coefficient and T2 relaxation during radiotherapy for prostate cancer. *Journal of Magnetic Resonance Imaging*, 37(4):909–916, 2013.

[21] Gang Cai, Ye Xu, Ji Zhu, Wei Lie Gu, Shuai Zhang, Xue Jun Ma, San Jun Cai, and Zhen Zhang. Diffusion-weighted magnetic resonance imaging for predicting the response of rectal cancer to neoadjuvant concurrent chemoradiation. *World Journal of Gastroenterology*, 19(33):5520–5527, 2013.

[22] Lan Wang, Lihong Liu, Chun Han, Shutang Liu, Hua Tian, Zhensheng Li, Xuejiao Ren, Gaofeng Shi, Qi Wang, and Guangda Wang. The diffusion-weighted magnetic resonance imaging (DWI) predicts the early response of esophageal squamous cell carcinoma to concurrent chemoradiotherapy. *Radiotherapy and Oncology*, 121(2):246–251, 2016.

[23] Ying Liu, Haoran Sun, Renju Bai, and Zhaoxiang Ye. Time-window of early detection of response to concurrent chemoradiation in cervical cancer by using diffusion-weighted MR imaging: A pilot study. *Radiation Oncology*, 10(1):1–8, 2015.

[24] Patrick A. Hein, Christian Kremser, Werner Judmaier, Jürgen Griebel, Karl Peter Pfeiffer, Alfons Kreczy, Eugen B. Hug, Peter Lukas, and Alexander F. DeVries. Diffusion-weighted magnetic resonance imaging for monitoring diffusion changes in rectal carcinoma during combined, preoperative chemoradiation: Preliminary results of a prospective study. *European Journal of Radiology*, 45(3):214–222, 2003.

[25] Chuang Wang, Andreas Rimner, Yu Chi Hu, Neelam Tyagi, Jue Jiang, Ellen Yorke, Sadegh Riyahi,

Gig Mageras, Joseph O. Deasy, and Pengpeng Zhang. Toward predicting the evolution of lung tumors during radiotherapy observed on a longitudinal MR imaging study via a deep learning algorithm. *Medical Physics*, 46(10):4699–4707, 2019.

[26] Fatemeh Nasiri and Oscar Acosta-Tamayo. Mixed-Effect Modeling for Longitudinal Prediction of Cancer Tumor. *arXiv*, pages 1–14, 2018.

[27] Sofie Rahbek, Kristoffer H. Madsen, Henrik Lundell, Faisal Mahmood, and Lars G. Hanson. Data-driven separation of MRI signal components for tissue characterization. *Journal of Magnetic Resonance*, 333(107103):1–11, 2021.

[28] Andreas M. Hötker, André Lollert, Yousef Mazaheri, Sabine Müller, Jens Peter Schenk, Philipp C. Mildenberger, Oguz Akin, Norbert Graf, and Gundula Staatz. Diffusion-weighted MRI in the assessment of nephroblastoma: results of a multi-center trial. *Abdominal Radiology*, 45(10):3202–3212, 2020.

[29] Mohammad Najafi, Hamid Soltanian-Zadeh, Kourosh Jafari-Khouzani, Lisa Scarpace, and Tom Mikkelsen. Prediction of glioblastoma multiform response to bevacizumab treatment using multi-parametric MRI. *PLoS ONE*, 7(1):1–11, 2012.

[30] Özlem Gürünlü Alma. Comparison of Robust Regression Methods in Linear Regression. *International Journal of Contemporary Mathematical Sciences*, 6(9):409–421, 2011.

[31] Faisal Mahmood, Helle H. Johannesen, Poul Geertsen, and Rasmus H. Hansen. Ultra-early apparent diffusion coefficient change indicates irradiation and predicts radiotherapy outcome in brain metastases. *Acta Oncologica*, 56(11):1651–1653, 2017.

[32] Bradford A. Moffat, Thomas L. Chenevert, Theodore S. Lawrence, Charles R. Meyer, Timothy D. Johnson, Qian Dong, Christina Tsien, Suresh Mukherji, Douglas J. Quint, Stephen S. Gebarski, Patricia L. Robertson, Larry R. Junck, Alnawaz Rehemtulla, and Brian D. Ross. Functional diffusion map: A noninvasive MRI biomarker for early stratification of clinical brain tumor response. *Proceedings of the National Academy of Sciences of the United States of America*, 102(15):5524–5529, 2005.

[33] David W Hosmer and Stanley Lemeshow. *Applied Logistic Regression*. John Wiley and Sons Inc., second edition, 2000.

[34] Zoran Bursac, Clinton Heath Gauss, David Keith Williams, and David W. Hosmer. Purposeful selection of variables in logistic regression. *Source Code for Biology and Medicine*, 3:1–8, 2008.

[35] Markus Ojala and Gemma C. Garriga. Permutation tests for studying classifier performance. *Journal of Machine Learning Research*, 11:1833–1863, 2010.

[36] Michal R. Tomaszewski, William Dominguez-Viqueira, Antonio Ortiz, Yu Shi, James R. Costello, Heiko Enderling, Stephen A. Rosenberg, and Robert J. Gillies. Heterogeneity analysis of MRI T2 maps for measurement of early tumor response to radiotherapy. *NMR in Biomedicine*, 34(3):1–12, 2021.

[37] Faisal Mahmood, Helle H. Johannesen, Poul Geertsen, and Rasmus H. Hansen. Repeated diffusion MRI reveals earliest time point for stratification of radiotherapy response in brain metastases. *Physics in Medicine and Biology*, 62(8):2990–3002, 2017.

[38] William Penny, Karl Friston, John Ashburner, Stefan Kiebel, and Thomas Nichols. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier science, 2007.

[39] Shigeyoshi Saito, Mana Tsugeno, Daichi Koto, Yuki Mori, Yoshichika Yoshioka, Satoshi Nohara, and Kenya Murase. Impact of surface coating and particle size on the uptake of small and ultrasmall superparamagnetic iron oxide nanoparticles by macrophages. *International Journal of Nanomedicine*, 7(May 2014):5415–5421, 2012.

[40] Rajamanickam Baskar, Jiawen Dai, Nei Wenlong, Richard Yeo, and Kheng Wei Yeoh. Biological response of cancer cells to radiation treatment. *Frontiers in Molecular Biosciences*, 1(NOV):1–9, 2014.

[41] David Eriksson and Torgny Stigbrand. Radiation-induced cell death mechanisms. *Tumor Biology*, 31(4):363–372, 2010.

[42] Daniel M Patterson, Anwar R Padhani, and David J Collins. Technology Insight: water diffusion

MRI—a potential new biomarker of response to cancer therapy. *Nature Clinical Practice Oncology*, 5(4):220–233, 2008.

[43] Ying Shi Sun, Yong Cui, Lei Tang, Li Ping Qi, Ning Wang, Xiao Yan Zhang, Kun Cao, and Xiao Peng Zhang. Early evaluation of cancer response by a new functional biomarker: Apparent diffusion coefficient. *American Journal of Roentgenology*, 197(1):23–29, 2011.

[44] E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verweij. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*, 45(2):228–247, 2009.

[45] Faisal Mahmood, Helle Hjorth, Poul Geertsen, and Rasmus Hvass. Diffusion MRI outlined viable tumour volume beats GTV in intra-treatment stratification of outcome. *Radiotherapy and Oncology*, 144:121–126, 2020.

[46] Doenja M.J. Lambregts, Geerard L. Beets, Monique Maas, Luís Curvo-Semedo, Alfons G.H. Kessels, Thomas Thywissen, and Regina G.H. Beets-Tan. Tumour ADC measurements in rectal cancer: Effect of ROI methods on ADC values and interobserver variability. *European Radiology*, 21(12):2567–2574, 2011.

[47] Masoumeh Gity, Behnaz Moradi, Rasool Arami, Ali Arabkheradmand, and Mohamad Ali Kazemi. Two different methods of region-of-interest placement for differentiation of benign and malignant breast lesions by apparent diffusion coefficient value. *Asian Pacific Journal of Cancer Prevention*, 19(10):2765–2770, 2018.

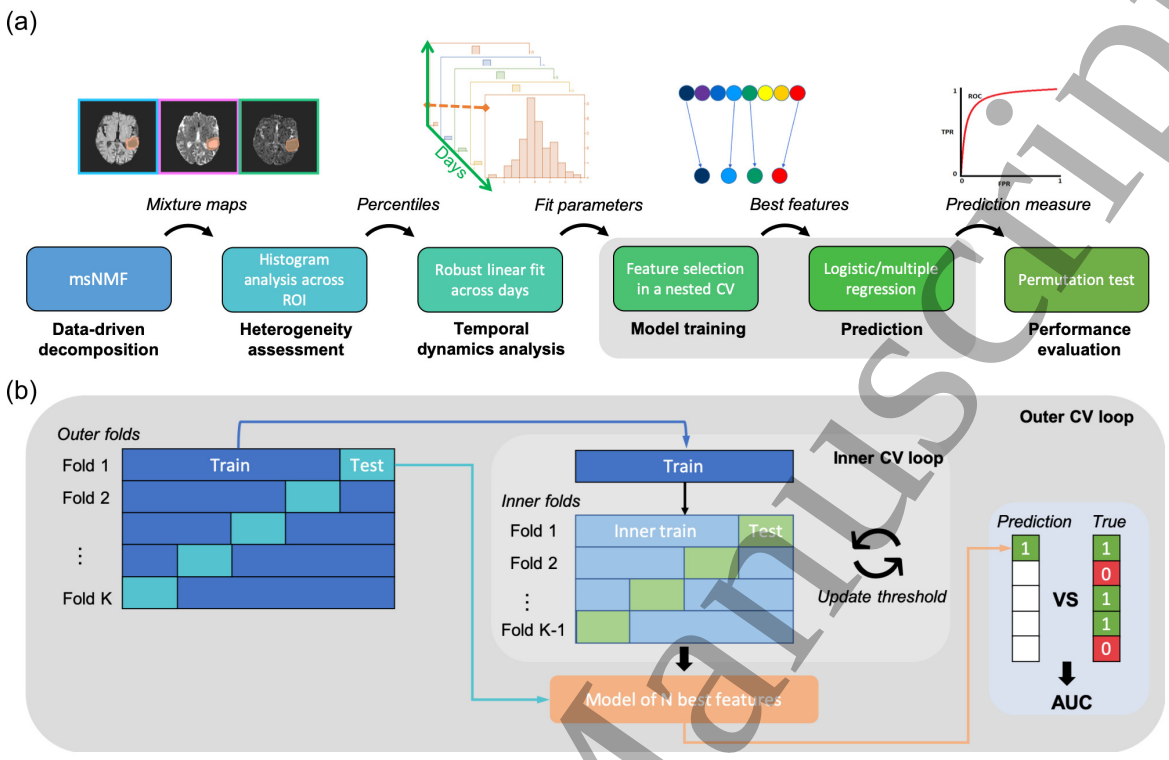

Figure 1. (a) The full prediction framework from the data decomposition to the final prediction. A nested cross-validation procedure was used for the feature selection and model training (gray box). This process is illustrated more detailed in (b).

FIGURES 20

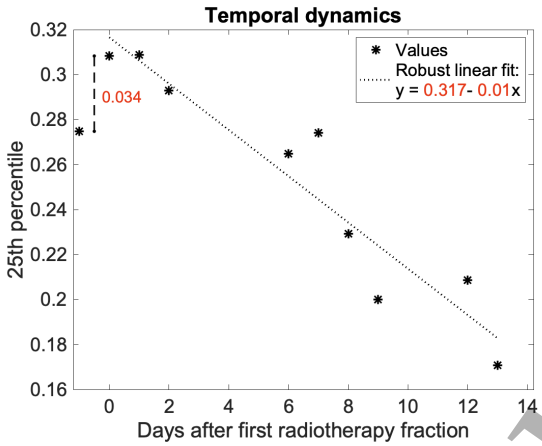

Figure 2. An example of the temporal fit using a Hubert loss function (dotted line). "Days after first treatment" is used as explanatory variable. The 25th percentile of component 1 for a random tumor is used as example data. The fitting parameters together with the difference to baseline (dashed line) are marked with a red font as these constitute the prediction features for this given case.

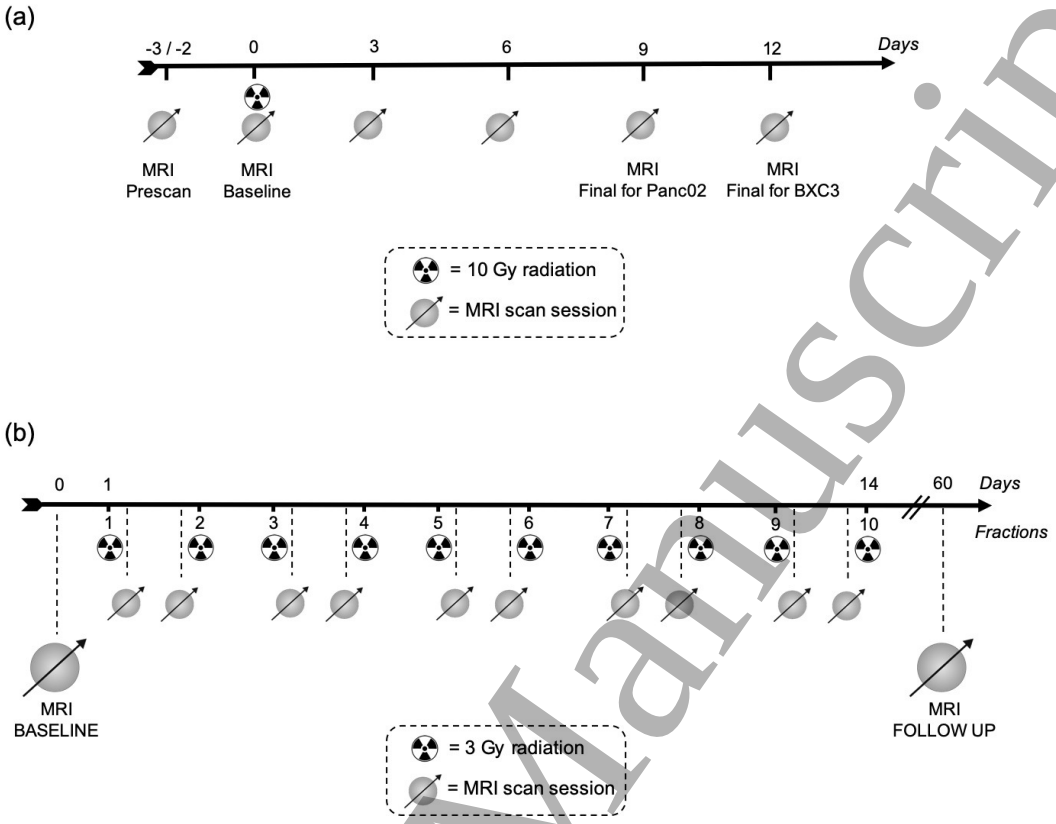

Figure 3. Timeline of the radiotherapy and MRI events, where the MRI baseline scan (last scan before radiation) defines "day 0". (a): The events for the mice pancreatic cancer study. (b): The events for the human brain metastases study. Notice the interleaved structure with an MRI scan right after and right before every second radiotherapy fraction*.

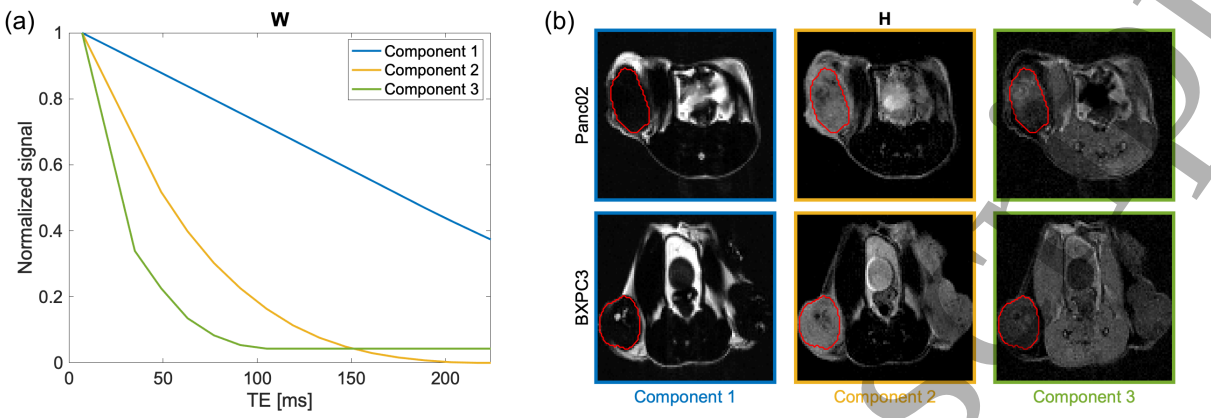

Figure 4. (a) The msNMF signal components (**W**) of mouse data. (b) The associated mixture maps (**H**) for two mice: one from each group. The red contour indicates the tumor ROI. The images are presented with a common, arbitrary intensity scale.

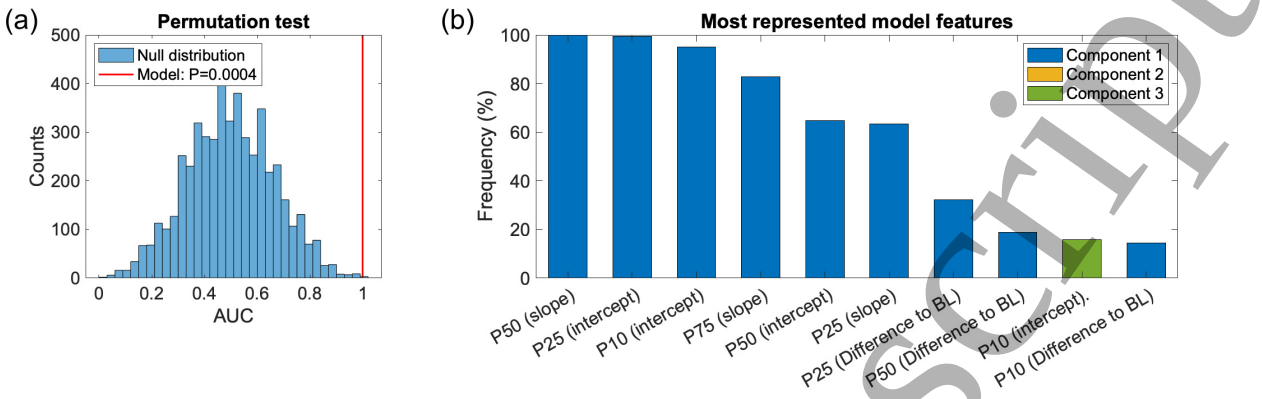

Figure 5. The result of the logistic regression for classification of tumor type. (a): The model AUC score (red line) compared to the null distribution generated from 5000 runs with permuted class labels (P=0.0004). (b): The most represented model features. The bar color indicates msNMF component, and PXX indicates percentile XX. The "slope" and "intercept" are from the temporal linear regression model. The "difference to BL" refers to the difference between the baseline scan and the subsequent scan at day 3.

FIGURES 24

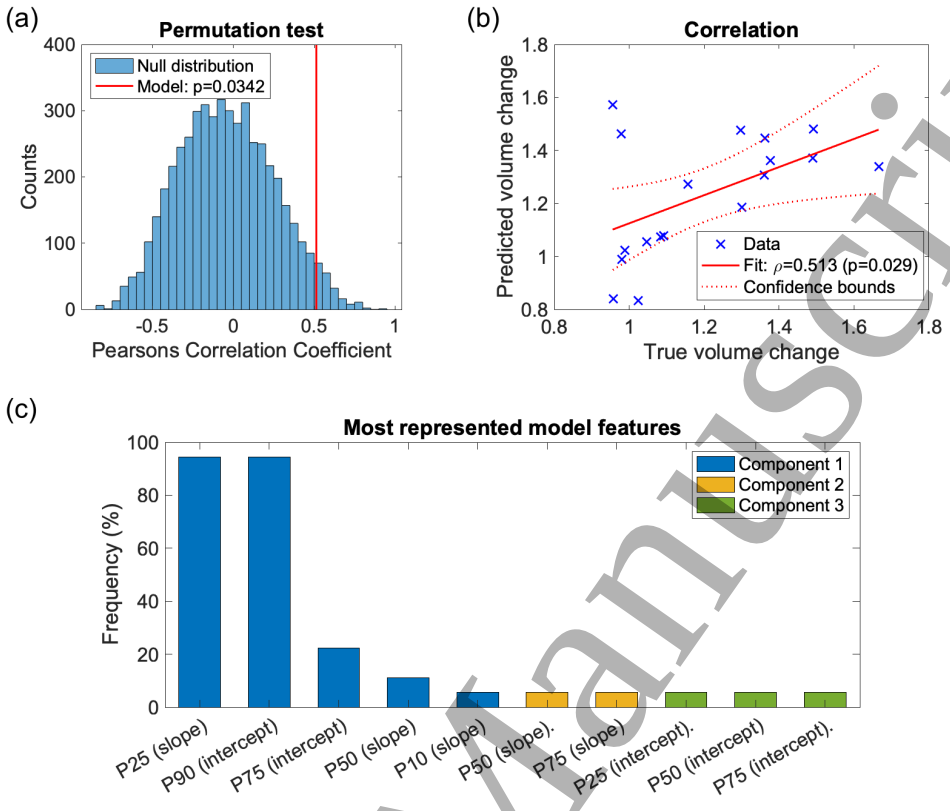

Figure 6. The result of the multiple linear regression for prediction of tumor volume change. (a): The model Pearson's correlation score (red line) compared to the null distribution generated from 5000 runs with permuted labels (P=0.034). (b): True versus predicted volume changes together with a linear fit presented with 95 % confidence bounds. (c): The most represented model features in the 18 CV runs.

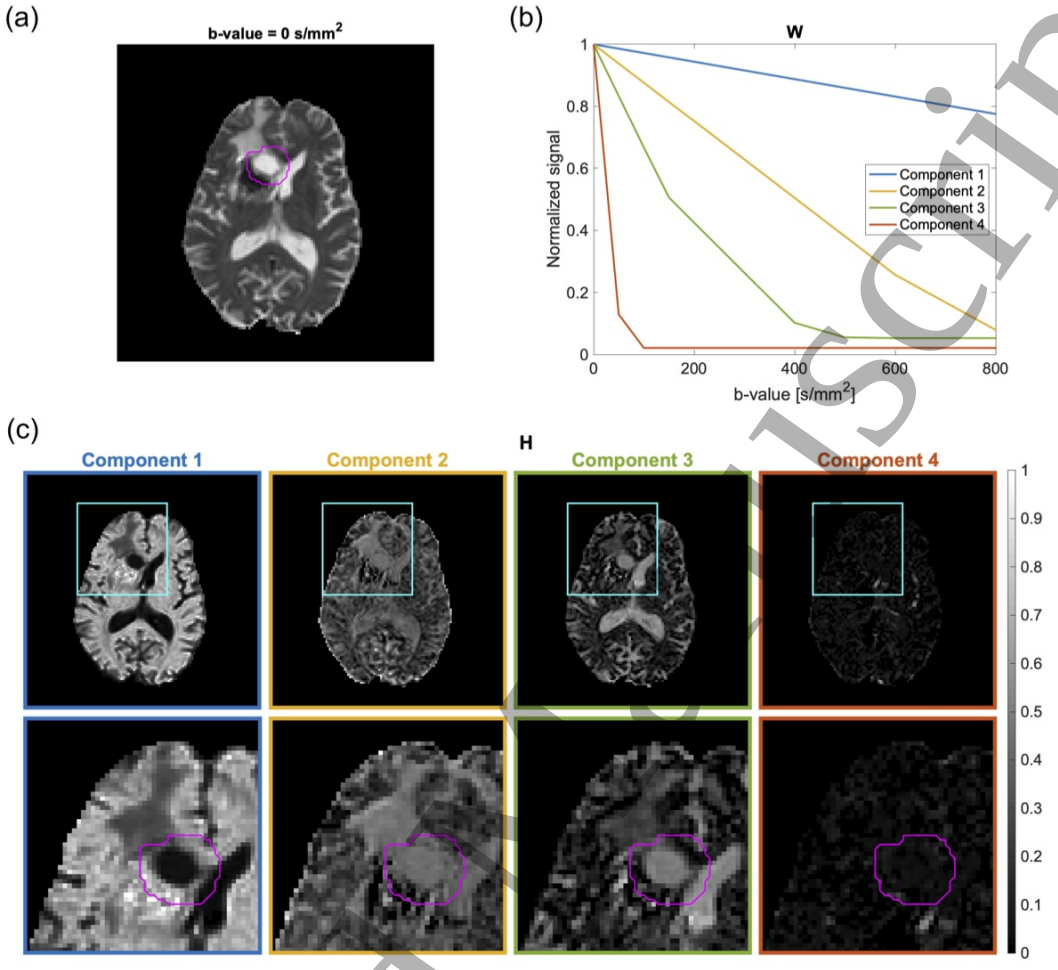

Figure 7. The msNMF components for an example patient with a brain metastasis, at scan day 1. (a): b-value=0 s/mm^2 image from the initial scan session. The magenta ROI is the result of dilating the radiologist's delineation. (b): The signal components (**W**). (c): The associated normalized mixture maps (**H**) indicated by frame colors. The cyan box surrounds the tumor area and marks the region magnified in the bottom row of images.

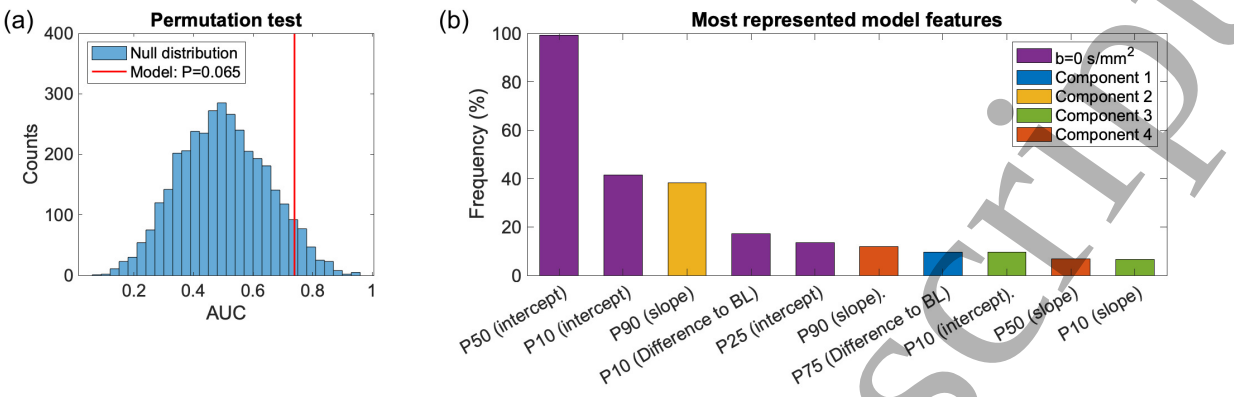

Figure 8. The result of the logistic regression for classification of metastasis into responders/non-responders. (a): The model AUC score (red line) compared to the null distribution generated from 5000 runs with permuted class labels (P=0.065). (b): The most represented model features. The bar color indicates signal/component, and PXX indicates percentile XX. The "slope" and "intercept" are from the temporal linear regression model. The "difference to BL" refers to the difference between the baseline scan and the subsequent scan (day 1).