Bone 120 (2019) 254–261

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

Prediction of early metastatic disease in experimental breast cancer bone metastasis by combining PET/CT and MRI parameters to a Model-Averaged Neural Network

Stephan Ellmann[a,*], Lisa Seyler[a], Jochen Evers[a], Henrik Heinen[a], Aline Bozec[b], Olaf Prante[c], Torsten Kuwert[c], Michael Uder[a], Tobias Bäuerle[a]

[a] Department of Radiology, Universitätsklinikum Erlangen, 91054 Erlangen, Germany
[b] Department of Internal Medicine 3, Rheumatology and Immunology, Universitätsklinikum Erlangen, 91054 Erlangen, Germany
[c] Department of Nuclear Medicine, Universitätsklinikum Erlangen, 91054 Erlangen, Germany

ARTICLE INFO

Keywords:
Bone metastases
Disseminated tumor cells
Breast cancer
Multiparametric imaging
Machine learning
Neural networks

ABSTRACT

Macrometastases in bone are preceded by bone marrow invasion of disseminated tumor cells. This study combined functional imaging parameters from FDG-PET/CT and MRI in a rat model of breast cancer bone metastases to a Model-averaged Neural Network (avNNet) for the detection of early metastatic disease and prediction of future macrometastases. Metastases were induced in 28 rats by injecting MDA-MB-231 breast cancer cells into the right superficial epigastric artery, resulting in the growth of osseous metastases in the right hind leg of the animals. All animals received FDG-PET/CT and MRI at days 0, 10, 20 and 30 after tumor cell injection. In total, 18/28 rats presented with metastases at days 20 or 30 (64.3%). None of the animals featured morphologic bone lesions during imaging at day 10, and the imaging parameters acquired at day 10 did not differ significantly between animals with metastases at or after day 20 and those without (all $p > 0.3$). The avNNet trained with the imaging parameters acquired at day 10, however, achieved an accuracy of 85.7% (95% CI 67.3–96.0%) in predicting future macrometastatic disease (ROC_{AUC} 0.90; 95% CI 0.76–1.00), and significantly outperformed the predictive capacities of all single parameters (all $p \leq 0.02$). The integration of functional FDG-PET/CT and MRI parameters into an avNNet can thus be used to predict macrometastatic disease with high accuracy, and their combination might serve as a surrogate marker for bone marrow invasion as an early metastatic process that is commonly missed during conventional staging examinations.

1. Introduction

Breast cancer is a leading malignancy in the western world and is known for its preference for the skeletal system in metastatic disease [1]. Affected patients experience a median survival time of 20 months when diagnosed with skeletal metastases [2]. Although the frequency of bone metastases is 5% at the time of primary diagnosis [3,4], bone metastases are found in one third of all patients with recurrent disease [5].

Metastatic skeletal lesions provide a challenge during staging, as they can only be captured with imaging techniques when sufficient bone loss has occurred [6]. Computed tomography (CT) offers higher sensitivity compared with conventional radiography, as it lacks the superposition of anatomical structures along with higher spatial resolution of morphological details [7,8]. Magnetic resonance imaging (MRI) offers a sensitivity and specificity of 70–100% for the detection of bony metastases. A meta-analysis of patients with bony metastases of breast cancer described a pooled sensitivity for MRI of 97%, which is significantly higher than for ^{18}F-fluorodeoxyglucose positron emission tomography (FDG-PET; 83%) and skeletal scintigraphy (87%) [8,9].

It is important to note that the aforementioned methods aim to detect metastases already present at a macroscopic level. However,

Abbreviations: CT, Computed Tomography; MRI, Magnetic Resonance Imaging; FDG-PET, Fluorodeoxyglucose Positron Emission Tomography; DTC, Disseminated Tumor Cells; avNNet, Model-Averaged Neural Network; SUV, Standardized Uptake Value; DCE, Dynamic Contrast Enhancement; ROC, Receiver Operating Characteristic; AUC, Area under the Curve; TTP, Time to Peak; PE, Peak Enhancement

* Corresponding author at: Department of Radiology, Universitätsklinikum Erlangen, Maximiliansplatz 3, 91054 Erlangen, Germany.
E-mail addresses: stephan.ellmann@uk-erlangen.de (S. Ellmann), lisa.seyler@uk-erlangen.de (L. Seyler), aline.bozec@uk-erlangen.de (A. Bozec), olaf.prante@uk-erlangen.de (O. Prante), torsten.kuwert@uk-erlangen.de (T. Kuwert), michael.uder@uk-erlangen.de (M. Uder), tobias.baeuerle@uk-erlangen.de (T. Bäuerle).

https://doi.org/10.1016/j.bone.2018.11.008
Received 24 July 2018; Received in revised form 7 November 2018; Accepted 12 November 2018
Available online 13 November 2018
8756-3282/ © 2018 Elsevier Inc. All rights reserved.

bone marrow invasion of tumor cells precedes macrometastasis and is considered as micrometastatic disease, in which disseminated tumor cells (DTC) reach the bone marrow and interact with the stroma cells that provide a fertile soil [10]. If the DTC are metabolically active and proliferate within the bone marrow, they will progress to form macrometastases. The initial bone marrow invasion including the interaction with the microenvironment can thus be considered an early step of metastases formation, which is currently missed during conventional staging examinations [11,12].

Though the information from different imaging modalities alone cannot sufficiently detect these early changes in bone marrow and thus not reliably predict the formation of macroscopic metastases, a combination of different imaging parameters might perform better. Recent investigations have focused on artificial neural networks to combine several parameters and perform predictive classification tasks with high accuracy [13]. In the human brain, nerve cells are some kind of nodes interconnected with each other. The connections that provide advantages in behaviour, body movement, memory or mental activity will become stronger over time, a process we refer to as learning. Artificial neural networks are combinations of artificial nodes organized into layers, with the predictor parameters as the bottom layer, and the output as the top layer. An additional intermediate layer contains hidden neuronal nodes which receive input from the predictors and form the output. The interconnections between those artificial neuronal nodes are initially weighted in a random fashion, but adapted during training of the network. This allows trained artificial networks to perform e.g. classification tasks with high accuracy. Their accuracy can be further increased by averaging the outputs of several models, resulting in Averaged-Model Neural Networks (avNNet) [14]. The underlying rationale is that the bias of such a network ensemble remains unchanged, while the variance error decreases considerably [14]. This is in contrast to other ensemble methods such as boosting that aim to reduce the bias of the combined estimator.

This study combines functional MRI and PET imaging parameters in an animal model of breast cancer bone metastases to an avNNet for the first time. Using appropriate input variables within this avNNet, the goal was to reliably predict in advance whether or not a given animal will develop osseous macrometastases when no visually perceivable metastases are present. The identified imaging markers could serve as potential surrogate markers of early metastatic disease.

2. Materials and methods

2.1. Cell line

The human estrogen-independent breast cancer cell line MDA-MB-231 was obtained from American type culture collection (ATCC). MDA-MB-231 cells were cultured in RPMI-1640 (Invitrogen, Karlsruhe, Germany) and supplemented with 10% fetal calf serum (Sigma, Taufkirchen, Germany). All cultures were kept under controlled conditions (37 °C, 5% CO_2, humidified atmosphere) and passaged 2–3 times a week.

2.2. Animal model of breast cancer bone metastases

Twenty-eight rats (RNU strain) were obtained from Charles River, Germany, and housed at the central animal facility of the University of Erlangen-Nuremberg. To avoid potential infections, rats were kept isolated and were fed with pathogen-free food. Clinical symptoms including body weight, relieving posture of the right hind leg, overall behavior, appearance and general health condition were monitored daily. All care and experimental procedures were performed in accordance with national and regional legislation on animal protection, and all animal procedures were approved by the State Government of Franconia, Germany (reference number 55.2 DMS-2532-2-228).

For metastases induction, four-weeks-old animals received isoflurane anesthesia (2%, 2 L/min) and a subcutaneous injection of carprofen (4 mg/kg) prior to surgery. An incision 2–3 cm in length was made in the right inguinal region of the animals. Subsequently, the arterial branches of the right hind leg were dissected. The flow of the femoral artery was temporarily occluded by clips that were placed proximal and distal of the superficial epigastric artery. The superficial epigastric artery was ligated distally, which enabled this vessel to be opened without causing bleeding. Following the insertion of a needle, 1.5×10^5 MDA-MB-231 breast cancer cells suspended in 0.2 mL media were slowly injected into the superficial epigastric artery and directed to the descending genicular and popliteal arteries by virtue of the clips [15].

As any other animal model, this model also features a tumor-take rate of < 100%, so that every experimental set-up would yield both animals that bear metastases and animals that do not. The aim of this study was to predict the development of macrometastases at an early time point, at which neither MR nor PET/CT imaging yielded differences between animals with and without development of metastases.

2.3. Imaging protocols

Imaging was performed prior to surgery, as well as at day 10, 20 and 30 post-surgery. Animals that did not develop metastases until day 30 were additionally imaged at day 45, whereas animals with metastases were sacrificed by cervical dislocation under isoflurane anesthesia (2%, 2 L/min) subsequent to the imaging procedure on day 30. Both FDG-PET/CT and MRI were performed within 24 h under isoflurane anesthesia (2%, 2 L/min). MRI was performed on a 7 Tesla MRI (ClinScan 70/30, Bruker, Ettlingen, Germany) using a dedicated body coil (Bruker) and the sequences listed in Table 1. For dynamic contrast enhancement measurements, 0.1 mmol gadobutrol (Gadovist, Bayer Schering, Germany) per kg body weight were injected into the tail vein of the rats after 5 of 100 dynamic cycles.

PET/CT was performed on a Siemens Inveon combined PET/CT scanner (Siemens Healthineers, Erlangen, Germany). Animals were fasted for 12 h prior to PET/CT examination, and then received an injection of 6 MBq ^{18}F-fluorodeoxyglucose (^{18}F-FDG) into the tail vein. To allow for sufficient distribution of the tracer, animals were kept under anesthesia for another 30 min and then subjected to a CT examination (duration 10 min, tube voltage 80 kV, tube current 500 μA, isotropic resolution 48.9 μm), followed by a static PET acquisition (duration 15 min, lower/upper discriminatory level 350/650 keV, timing window 3.438 ns). The reconstructed PET images were converted to SUV images as described below.

Table 1
MRI acquisition parameters.

Parameter	Sequence		
	T2w TSE	DCE FLASH	T1w TSE
TR	8654 ms	3.9 ms	795 ms
TE	37 ms	0.88 ms	9 ms
ETL	7	n/a	1
Flip angle	180°	25°	90°
Field of view	65 × 55 mm^2	65 × 54 mm^2	65 × 54 mm^2
Matrix	320 × 272	256 × 216	320 × 270
Slice thickness	1 mm	1 mm	1 mm
Number of slices	60	8 × 100 time points	60
Averages	1	1	2
Duration [min:sec]	11:24	8:25	6:18

Acquisition parameters are given for the sequences: T2w turbo spin echo (TSE), Dynamic contrast enhancement fast low angle shot (DCE FLASH), and T1w TSE, along with their repetition times (TR), echo times (TE), echo train lengths (ETL), flip angles, field of view, matrix size, slice thickness, the number of acquired slices, averages, and sequence duration.

2.4. Image analysis

MR and CT image analysis was performed using Osirix [16] and the freeware dynamic contrast enhancement (DCE) tool [17]. The PET images were analyzed using Inveon Acquisition Software (Siemens Healthineers, Erlangen, Germany). The metastases were identified in consensus by a board-approved radiologist with 15 years of experience in oncologic imaging (T.B.) and a radiology resident with 5 years of experience (S.E.). The visible presence of metastases, determined by cortical destruction in CT, T2w hyperintense or T1w contrast-enhancing lesions in MRI, served as a reference standard.

In the axial MR images acquired at day 10, circular 2-dimensional regions-of-interest (ROI) with a target size of 1.5 mm^2 were placed in the proximal tibial shaft's bone marrow of the right hind leg (performed by S.E) and subjected to DCE analyses to determine the time to peak (TTP; *se*conds), peak enhancement (PE; arbitrary units), area under the curve (AUC; arbitrary units), and contrast media washout (percent). The PE and AUC values were normalized by dividing the measured values by the respective measurements acquired in the back muscles of the animal.

For the PET analysis, mean and maximum activities in the proximal tibial shaft of the right hind leg were determined using oval 2-dimensional ROI in coronal orientation (target size 10 mm^2). Mean and maximum standardized uptake values (SUV_{mean} and SUV_{max}, respectively) were calculated by dividing the obtained mean and maximum activities by the injected activity, and multiplying the result with the weight of the individual animal.

At the time of analysis, the reader was not aware whether or not a given animal would develop macrometastases.

2.5. Statistical analyses and development of a Neural Network

All statistical analyses were performed using GraphPad Prism 6 (GraphPad Software, La Jolla, CA, USA) and RStudio 3.4.1 [18]. For intergroup comparisons, Wilcoxon rank sum tests were applied to assess the significance of differences. To assess the predictive abilities of every single measured parameter and the resulting avNNet, the areas under the curves of Receiver Operating Characteristics (ROC_{AUC}) were calculated, and subsequently compared using DeLong tests. Linear correlations were assessed using the Pearson correlation method. For all statistical tests, a significance level (alpha) of 0.05 was chosen, and p-values < 0.05 were considered statistically significant.

avNNet development was performed with Rstudio [18] and additional packages including caret [19], nnet [20] and NeuralNetTools [21]. The aim was to create an avNNet that predicts the presence of macrometastases at day 20 or 30 based on the input variables acquired at day 10 post-surgery.

The ReliefF algorithm was used as a filter method to preselect relevant features for the avNNet [22,23]. The features selected by ReliefF were kept as predictors for the final model, while the irrelevant features were omitted. Therefore, ReliefF determines the features' weights regarding categorical classification tasks, with the algorithm penalizing the features that assign different values to neighbors of the same class, and rewarding features that assign different values to neighbors of different classes.

The avNNet applied for the purpose of this study consisted of one intermediate layer containing hidden neurons as nodes, which receive input from the predictors and form the output, and consisted of five individual models with their output being averaged. Neural networks' performances are influenced by hyperparameters, including the number of hidden neurons and a variable referred to as decay, which restricts the weights of the neuronal connections. To determine the optimal combination of hidden neurons and the decay variable, grid searches were applied. The resulting optimal combination was then used to develop and train the model in the remainder of this study. Detailed additional avNNet parameters are presented in Table 2. Those parameters were used with their standard setting in the grid search as well as in the final training process. A schematic presentation of the avNNet architecture is presented in Supplementary Fig. 1.

Table 2
Neural network – detailed parameters.

Parameter	Setting
Regularization	L2
Activation function	Logistic
Loss function	Mean Squared Error (MSE)
Optimizer	Broyden–Fletcher–Goldfarb–Shanno (BFGS)
Stopping criteria	• Absolute tolerance = 1×10^{-4} • Relative tolerance = 1×10^{-8} • Maximum number of iterations = 100

To accurately assess the model's performance in predicting unknown data, a leave-one-out cross validation approach was used. Thus, the first animal's data was removed from the data set and the avNNet was trained using the data of the remaining 27 animals and subsequently used to predict the outcome of the formerly withheld animal. This process was iteratively repeated for every animal, so that each animal's outcome was predicted with data not part of the training procedure. Before being fed into the avNNet, the predictors underwent BoxCox transformation. This transformation was performed for every cross validation step to avoid data leakage.

3. Results

3.1. Development of metastases

At day 30, 18 of the 28 animals exhibited metastases in their right hind leg (tumor take rate 64.3%). Of these 18, metastases were already detectable at day 20 in 14 of them; in the remaining 4, metastases could be identified at day 30 post-surgery as T2w hyperintense, gadolinium-enhancing lesions in MRI and cortical destruction in CT (compare Figs. 1, 2). Ten animals did not develop metastases until day 30, and were observed until day 45 without detectable metastases. None of the 28 animals featured macroscopically visible suspicious lesions at day 10.

The first row depicts T2w MR images, the second row depicts T1w

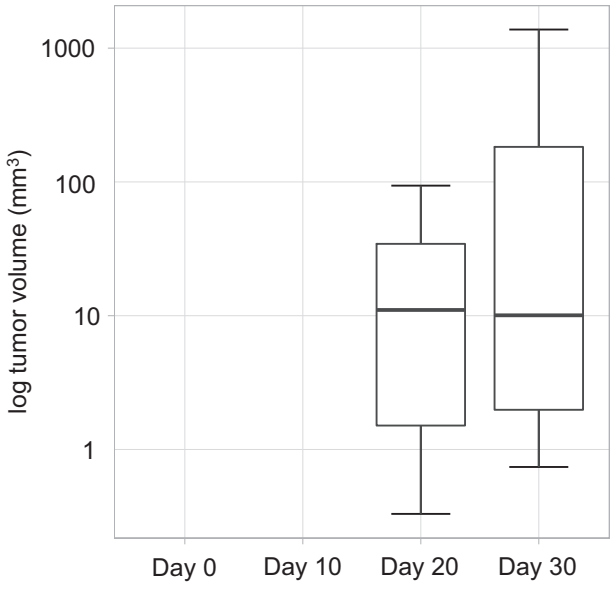

Fig. 1. Boxplot depicting the tumor volumes measured in MRI at day 0, 10, 20 and 30 after induction of metastases (n = 18 animals with metastases). None of the animals featured visible macrometastases before day 20.

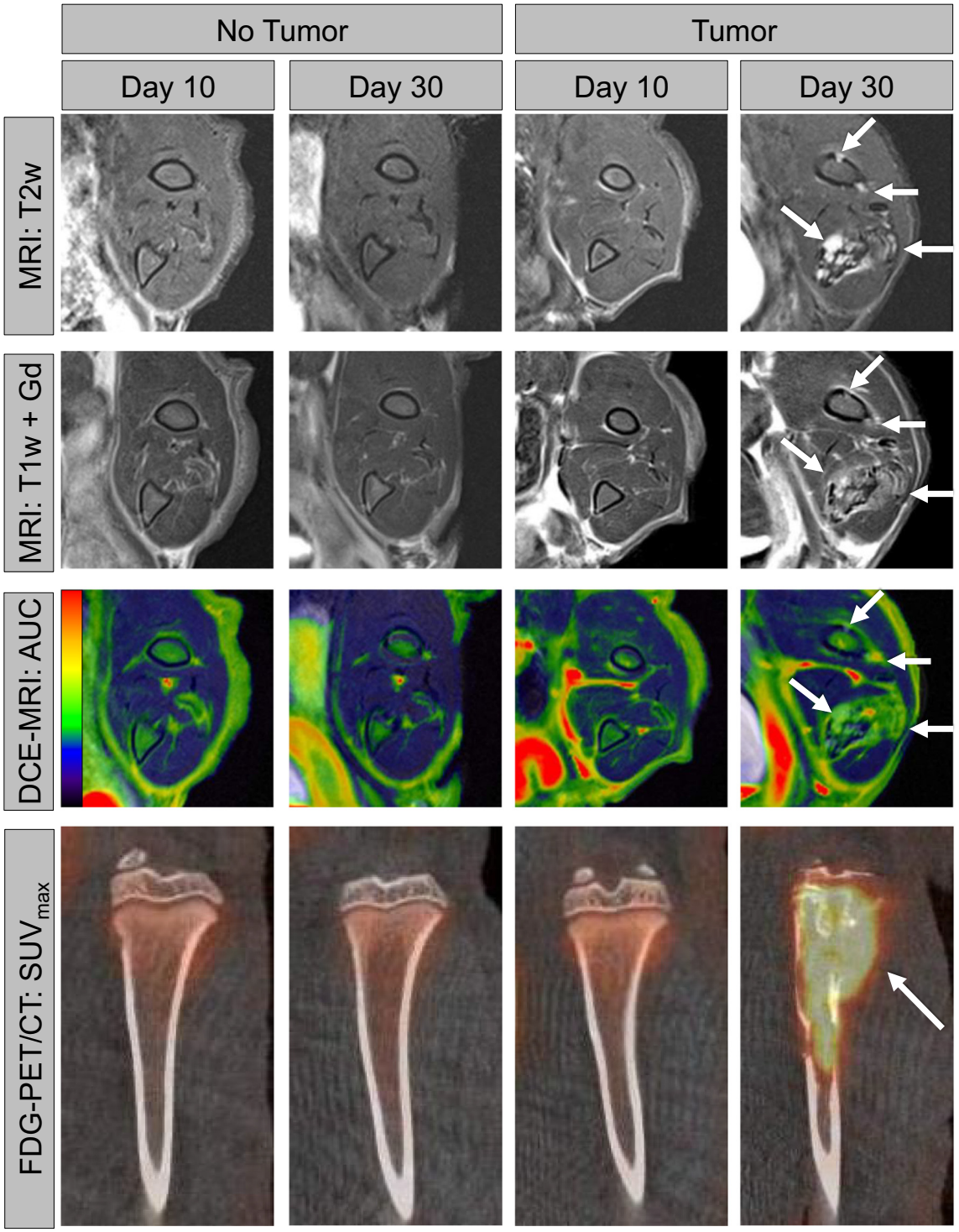

Fig. 2. Representative images of an animal's right hind leg that developed no metastases over the course of the study (two leftmost columns, with images from day 10 and day 30 presented) and an animal that developed metastases between day 10 and 30 (two rightmost columns, metastases marked with arrows).

post-contrast (Gd) MR images, and the third row depicts maps visualizing the area under the curve from the dynamic contrast enhancement (DCE) sequence. The fourth row presents the visualized maximum standardized uptake values (SUV_{max}) from FDG-PET fused with CT images.

3.2. Feature analysis and selection

Two features were acquired from FDG-PET imaging within the bone marrow of the proximal tibial shaft: SUV_{mean} and SUV_{max}. DCE MRI yielded the four features AUC, PE, TTP, and contrast media washout. At day 10, none of those six imaging features differed significantly between animals bearing a tumor at day 20 or 30 and animals that did not develop metastases during the course of the study (all $p > 0.3$; compare Fig. 3).

ROC analyses were performed to calculate the ROC_{AUC} and assess the predictive abilities of the acquired features in terms of future growth of metastases. Regarding PET parameters, SUV_{max} performed

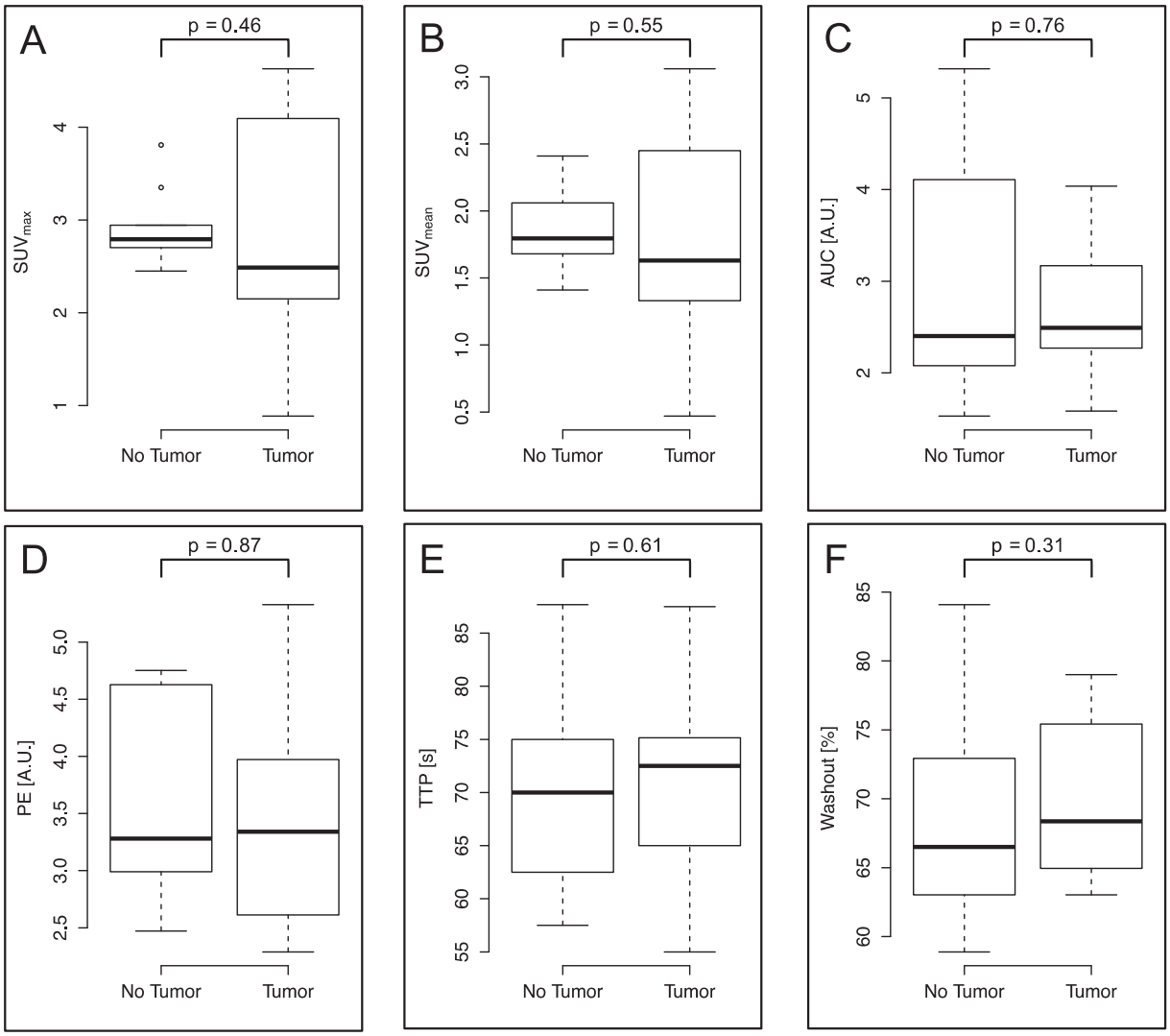

Fig. 3. Boxplots comparing the features measured at day 10 in FDG-PET/CT [A, B] and MRI [C–F], grouped by the reference standard of visible tumors at or after day 20. [A] Maximum standardized uptake value SUV_{max}; [B] mean standardized uptake value SUV_{mean}; [C] area under the curve AUC; [D] peak enhancement PE; [E] time to peak TTP; [F] contrast media washout. Boxplots follow the Tukey definition. Outliers are marked with circles. *p*-values were determined by Wilcoxon rank sum tests and are given above the plots.

better than SUV_{mean} (ROC_{AUC} 0.59 vs 0.57, compare Fig. 4). Among the DCE features, relative washout (ROC_{AUC} 0.62) performed better than TTP (ROC_{AUC} 0.56), AUC (ROC_{AUC} 0.54), and PE (ROC_{AUC} 0.48, compare Fig. 4).

3.3. Neural Network training and performance

SUV_{mean} was rejected as a predictor for the final model as it was highly correlated to SUV_{max} ($r = 0.95$, $p = 4.6 \times 10^{-15}$), but featured a lower ROC_{AUC}. From the remaining features, SUV_{max}, AUC and PE were selected as predictors for the final model. The initial grid search yielded 3 hidden neurons and a decay of 0.0005 as the optimal parameter combination for the subsequent training procedure (Fig. 5).

The training and leave-one-out cross validation procedure resulted in an overall accuracy of 85.7% (95% CI 67.3–96.0%) and an ROC_{AUC} of 0.90 (95% CI 0.76–1.00; Fig. 6a). The avNNet significantly outperformed the single predictors SUV_{max} ($ROC_{AUC} = 0.59$; $p = 0.02$), AUC ($ROC_{AUC} = 0.54$; $p = 0.02$), and PE ($ROC_{AUC} = 0.48$; $p = 0.003$).

Table 3 provides detailed diagnostic accuracy measures of the avNNet, including sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, ROC_{AUC} and overall accuracy, along with 95% confidence intervals, as well as the *p*-value for a comparison of the model's accuracy with the accuracy based on pure guessing ("No Information Rate").

In addition to its dichotomous output, the avNNet returned probabilistic values, as presented in Fig. 6B. For animals with future metastases, the avNNet returned a median probability of 0.960, whereas for animals without metastases it returned a median probability of 0.156 ($p = 0.0002$). A schematic illustration of the avNNet's architecture is presented in Supplementary Fig. 1.

4. Discussion

This study presents a novel application of avNNet to predict the development of macroscopic skeletal metastases in an experimental model of breast cancer bone metastases.

In this study, no visually perceivable differences were present 10 days after metastasis induction in PET/CT and MR imaging, and the quantified functional imaging features acquired at day 10 did not differ significantly between animals bearing macrometastases at day 20 or 30 and animals without metastases. The presented model, however, exhibited 85.7% accuracy, with 90% specificity and 83.3% sensitivity, in predicting growth of macrometastases, based on functional imaging predictors acquired at day 10. These results are derived from a

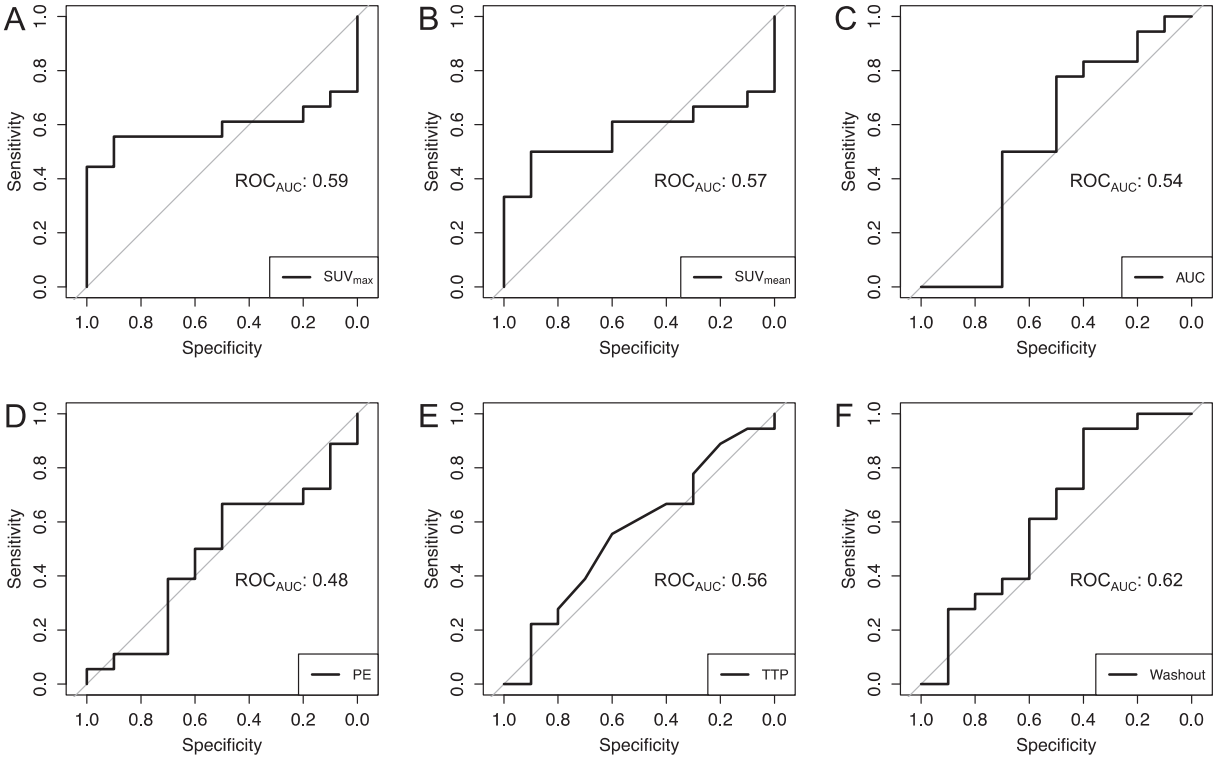

Fig. 4. Receiver Operating Characteristic (ROC) plots for the features acquired at day 10: [A] maximum standardized uptake value SUV_{max}; [B] mean standardized uptake value SUV_{mean}; [C] area under the curve AUC; [D] peak Enhancement PE; [E] time to peak TTP; [F] contrast media washout, along with the area under the curve values (ROC_{AUC}).

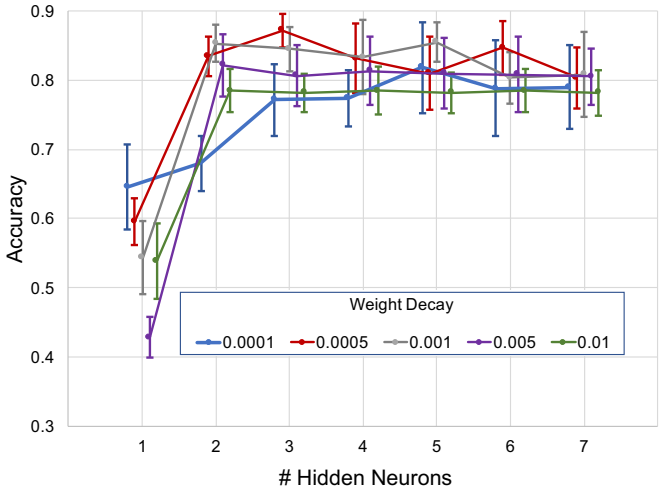

Fig. 5. Averaged results from 10 grid searches among different combinations of hidden neurons in the network's intermediate layer and weight decays for the avNNet. Highest accuracy was achieved with 3 hidden neurons and a weight decay of 0.0005. Error bars indicate standard deviations.

machine-learning based combination of functional imaging measures, with FDG-PET/CT-derived parameters reflecting bone marrow glucose metabolism, while parameters from the DCE-MRI represent microvascular perfusion of the bone marrow [24,25]. In this regard, the applied predictors appear to complement each other, since SUV_{max} (as an individual parameter) offered reasonable specificity while being quite insensitive, while PE and especially AUC provided higher sensitivity with mediocre specificity (Fig. 6A).

The potential independency of predictors that reflect perfusion and glucose metabolism, with the ability to obtain complementary information from a combination, has been described for malignant entities, including colorectal liver metastases and head-and-neck tumors [25–27]. However, this study, for the first time, applied this combination to assess bone marrow changes due to interaction between tumor cells and their microenvironment to predict macroscopic metastases.

An important concept regarding the pathogenesis of bony metastases is DTC. In the pathogenesis of breast cancer metastasis, these cells are known to extravasate into the bone marrow, interact with the microenvironment and proliferate influencing the outcome of patients [28,29]. DTC are supposed to be delivered to the bone marrow via the bloodstream [30]. With respect to the applied experimental model, an injection of tumor cells into the right superficial epigastric artery leads to site-specific metastases around the knee, with no other distant metastases ever observed [15], which can be considered to mimic the site-specific sowing of DTC. From a clinical perspective, it remains unclear whether these DTC in the bone marrow represent the population of cells to develop into metastatic foci, the systemic burden of micrometastatic disease, or a combination of both [31]. However, they entirely escape conventional staging examinations [11,12]. Locally proliferating DTC, however, seem to induce changes in microvascularization and glucose metabolism within the affected bone marrow-parameters that can be assessed with multiparameteric imaging approaches.

While the detection of altered bone marrow FDG-uptake has been investigated intensively for hematolymphoid disorders [32–34], few studies have examined the bone marrow infiltration of solid tumors. However, an ^{18}F-FDG PET/CT positive bone marrow signal could be determined as an indicator of early metastasis for solid tumors as well [35,36], with SUV_{max} offering high specificity [36] in accordance with this study's results. Apart from few retrospective studies utilizing FDG-PET/CT for the identification of microinvasion [35,36], imaging for the purpose of solid tumor staging has largely aimed at detecting metastases already present at a macroscopic level.

This study presents, for the first time, a machine-learning based approach for the detection of early metastatic disease. Nevertheless,

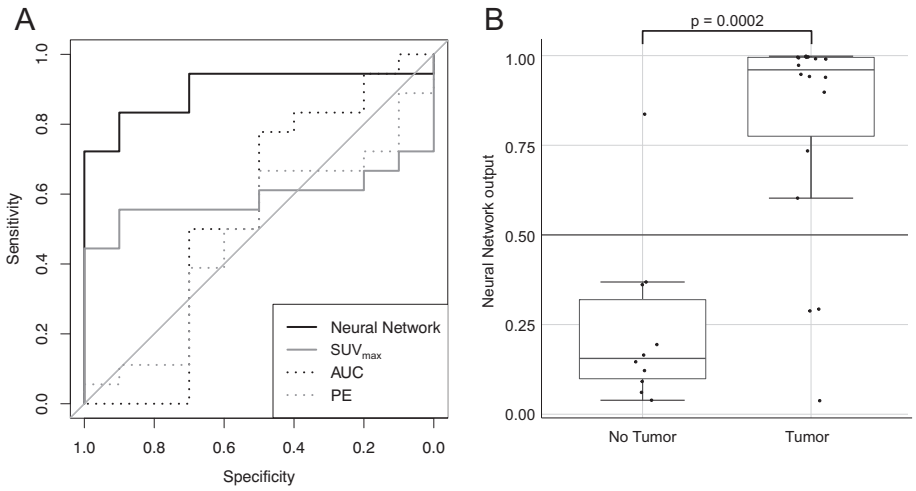

Fig. 6. [A] Receiver Operating Characteristic plot for the Model-Averaged Neural Network (black solid line) and the single predictors: maximum standardized uptake value SUV_{max} (gray solid line), area under the curve AUC (black dotted line), and peak enhancement PE (gray dotted line).
[B] Boxplot depicting the probabilistic results of the Model-Averaged Neural Network. For animals developing metastases, the median probability of the Model-Averaged Neural Network was 0.96, for animals without metastases 0.156 (Wilcoxon rank sum $p = 0.0002$).

Table 3
Diagnostic accuracy measures of the Model-Averaged Neural Network along with their 95% confidence intervals.

Measure	Estimation	95% Confidence Interval
Sensitivity	83.3%	58.6–96.4%
Specificity	90.0%	55.5–99.7%
Positive predictive value	93.8%	69.8–99.8%
Negative predictive value	75.0%	42.8–94.5%
Likelihood ratio of a positive test	8.33	1.28–54.11
Likelihood ratio of a negative test	0.19	0.07–0.53
ROC - area under the curve	0.90	0.76–1.00
Overall accuracy	85.7%	67.3–96.0%
p-Value (Accuracy vs. No Information Rate)	0.01	n.a.

there are several limitations:

The developed model relied on 28 individual animals, which is a rather low number for the purposes of machine learning. Ideally, one would randomly remove a fraction of animals from the entire data set and treat them as a test set, while training the avNNet on the remainder of the animals to subsequently predict the outcome of the test set. This procedure would, however, need considerably higher numbers of animals. To create a generalizable model, a leave-one-out cross validation was used in this study. This approach is common for small sample sizes and provides an almost unbiased estimate of a model's true generalization ability [37], but leads to imbalanced classes due to the 64.3% prevalence of macrometastatic disease. Moreover, the predictors of this model were preselected using ReliefF. Although this principle has been applied in several published studies [38–40], concerns have been raised over the possibility of too optimistic results due to feature preselection [41], especially when the number of predictor variables exceeds the number of training samples by far [42], e.g. in DNA microarray studies. Though this is not the case in the current study, a preselection bias cannot be totally excluded. This further endorses the need for a validation based on an independent test set, which is planned for future studies.

Second, this study relied on an animal model, and is thus considered experimental. As a result of the utilized site-specific model, it was clear that animals will – depending on the tumor take rate – either develop metastases around their right knee or not, but certainly nowhere else, providing an exceptionally defined area of interest to analyze. Predicting the development of macrometastases will be markedly more complicated when assessing the entire skeletal system to judge whether metastases will develop without the a priori knowledge of a specific site.

Third, this study presents a method to assess 10–20 days in advance whether macrometastatic disease will occur. The exact time window that allows this kind of prediction in advance remains unclear. Although, determining the maximum size of this window was not the aim of this study, it will be of particular interest for further studies.

5. Conclusions

This study describes the combination of functional imaging parameters to an avNNet to predict the development of macroscopic bone metastases, for the first time. Thus, the combination of PET/CT and MR imaging predictors might serve as a non-invasive surrogate marker for early metastatic disease. This study can therefore lay a foundation for the detection of bone metastasis at an earlier, microscopic state, by integrating functional and metabolic parameters.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bone.2018.11.008.

Funding

This work was supported by the German Research Foundation (DFG, Clinical Research Collaboration CRC 1181, subproject Z02; Priority Programme µBone, projects BA 4027/10-1 and BO 3811) and by the Emerging Fields Initiative (EFI) “Big Thera” of the Friedrich-Alexander-University Erlangen-Nuremberg.

Conflicts of interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

CRediT authorship contribution statement

Stephan Ellmann: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing-Original Draft, Writing-Review & Editing, Visualization. **Lisa Seyler:** Formal analysis, Investigation, Data curation, Writing-Review & Editing, Visualization. **Jochen Evers:** Formal analysis, Investigation, Writing-Review & Editing. **Henrik Heinen:** Formal analysis, Investigation, Writing-Review & Editing. **Aline Bozec:** Conceptualization, Methodology, Validation, Resources, Writing-Review & Editing, Project administration, Funding acquisition. **Olaf Prante:** Methodology, Validation, Resources, Writing-Review & Editing. **Torsten Kuwert:** Validation, Resources, Writing-Review & Editing, Funding acquisition. **Michael Uder:** Validation, Resources, Writing-Review & Editing, Funding acquisition. **Tobias Bäuerle:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Data curation, Writing-Review & Editing, Supervision, Project administration, Funding acquisition.

References

[1] K. Manders, L.V. van de Poll-Franse, G.-J. Creemers, G. Vreugdenhil, M.J.C. van der Sangen, G.A.P. Nieuwenhuijzen, R.M.H. Roumen, A.C. Voogd, Clinical management of women with metastatic breast cancer: a descriptive study according to age group, BMC Cancer 6 (2006) 179, , https://doi.org/10.1186/1471-2407-6-179.

[2] R.E. Coleman, R.D. Rubens, The clinical course of bone metastases from breast cancer, Br. J. Cancer 55 (1987) 61–66.

[3] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA Cancer J. Clin. 61 (2011) 69–90, https://doi.org/10.3322/caac.20107.

[4] A.Ø. Jensen, J.B. Jacobsen, M. Nørgaard, M. Yong, J.P. Fryzek, H.T. Sørensen, Incidence of bone metastases and skeletal-related events in breast cancer patients: a population-based cohort study in Denmark, BMC Cancer 11 (2011) 29, , https://doi.org/10.1186/1471-2407-11-29.

[5] I.H. Kunkler, M.V. Merrick, The value of non-staging skeletal scintigraphy in breast cancer, Clin. Radiol. 37 (1986) 561–562.

[6] T. Bäuerle, W. Semmler, Imaging response to systemic therapy for bone metastases, Eur. Radiol. 19 (2009) 2495–2507, https://doi.org/10.1007/s00330-009-1443-1.

[7] M. Sundaram, M.H. McGuire, Computed tomography or magnetic resonance for evaluating the solitary tumor or tumor-like lesion of bone? Skelet. Radiol. 17 (1988) 393–401.

[8] S. Ellmann, M. Beck, T. Kuwert, M. Uder, T. Bäuerle, Multimodal imaging of bone metastases: from preclinical to clinical applications, J. Orthop. Transl. 3 (2015) 166–177, https://doi.org/10.1016/J.JOT.2015.07.004.

[9] T. Liu, T. Cheng, W. Xu, W.-L. Yan, J. Liu, H.-L. Yang, A meta-analysis of 18FDG-PET, MRI and bone scintigraphy for diagnosis of bone metastases in patients with breast cancer, Skelet. Radiol. 40 (2011) 523–531, https://doi.org/10.1007/s00256-010-0963-8.

[10] S. D'Oronzo, J. Brown, R. Coleman, The role of biomarkers in the management of bone-homing malignancies, J. Bone Oncol. 9 (2017) 1–9, https://doi.org/10.1016/j.jbo.2017.09.001.

[11] S. Braun, K. Pantel, Clinical significance of occult metastatic cells in bone marrow of breast cancer patients, Oncologist 6 (2001) 125–132.

[12] S. Braun, R. Rosenberg, S. Thorban, N. Harbeck, Implications of occult metastatic cells for systemic cancer treatment in patients with breast or gastrointestinal cancer, Semin. Surg. Oncol. 20 (2001) 334–346.

[13] J.L. Patel, R.K. Goyal, Applications of artificial neural networks in medical science, Curr. Clin. Pharmacol. 2 (2007) 217–226.

[14] U. Naftaly, N. Intrator, D. Horn, Optimal ensemble averaging of neural networks, Netw. Comput. Neural Syst. 8 (1997) 283–296, https://doi.org/10.1088/0954-898X_8_3_004.

[15] T. Bäuerle, H. Adwan, F. Kiessling, H. Hilbig, F.P. Armbruster, M.R. Berger, Characterization of a rat model with site-specific bone metastasis induced by MDA-MB-231 breast cancer cells and its application to the effects of an antibody against bone sialoprotein, Int. J. Cancer 115 (2005) 177–186, https://doi.org/10.1002/ijc.20840.

[16] A. Rosset, L. Spadola, O. Ratib, OsiriX: an open-source software for navigating in multidimensional DICOM images, J. Digit. Imaging 17 (2004) 205–216, https://doi.org/10.1007/s10278-004-1014-6.

[17] K. Sung, DCE Tool Plugin, http://kyungs.bol.ucla.edu/software/DCE_tool/DCE_tool.html , Accessed date: 1 June 2018(n.d.).

[18] RStudio – Open Source and Enterprise-ready Professional Software for R, https://www.rstudio.com/, (2015) , Accessed date: 15 March 2017.

[19] M. Kuhn, CRAN - Package Caret, https://cran.r-project.org/web/packages/caret/index.html, (2016) , Accessed date: 15 March 2017.

[20] B. Ripley, Feed-Forward Neural Networks and Multinomial Log-Linear Models, [R package nnet version 7.3-12] (2009).

[21] M.W. Beck, NeuralNetTools: visualization and analysis tools for neural networks, J. Stat. Softw. 85 (2018) 1–20, https://doi.org/10.18637/jss.v085.i11.

[22] I. Kononenko, E. Šimec, M. Robnik-Šikonja, Overcoming the myopia of inductive learning algorithms with RELIEFF, Appl. Intell. 7 (1997) 39–55, https://doi.org/10.1023/A:1008280620621.

[23] Z.M. Hira, D.F. Gillies, A review of feature selection and feature extraction methods applied on microarray data, Adv. Bioinforma. 2015 (2015) 1–13, https://doi.org/10.1155/2015/198363.

[24] U. Tateishi, M. Miyake, T. Nagaoka, T. Terauchi, K. Kubota, T. Kinoshita, H. Daisaki, H.A. Macapinlac, Neoadjuvant chemotherapy in breast cancer: prediction of pathologic response with PET/CT and dynamic contrast-enhanced MR imaging—prospective assessment, Radiology 263 (2012) 53–63, https://doi.org/10.1148/radiol.12111177.

[25] S. De Bruyne, N. Van Damme, P. Smeets, L. Ferdinande, W. Ceelen, J. Mertens, C. Van de Wiele, R. Troisi, L. Libbrecht, S. Laurent, K. Geboes, M. Peeters, Value of DCE-MRI and FDG-PET/CT in the prediction of response to preoperative chemotherapy with bevacizumab for colorectal liver metastases, Br. J. Cancer 106 (2012) 1926–1933, https://doi.org/10.1038/bjc.2012.184.

[26] D. Vriens, H.W.M. van Laarhoven, J.J.A. van Asten, P.F.M. Krabbe, E.P. Visser, A. Heerschap, C.J.A. Punt, L.-F. de Geus-Oei, W.J.G. Oyen, Chemotherapy response monitoring of colorectal liver metastases by dynamic Gd-DTPA-enhanced MRI perfusion parameters and 18F-FDG PET metabolic rate, J. Nucl. Med. 50 (2009) 1777–1784, https://doi.org/10.2967/jnumed.109.064790.

[27] S. Bisdas, O. Seitz, M. Middendorp, N. Chambron-Pinho, T. Bisdas, T.J. Vogl, R. Hammerstingl, U. Ernemann, M.G. Mack, An exploratory pilot study into the association between microcirculatory parameters derived by MRI-based pharmacokinetic analysis and glucose utilization estimated by PET-CT imaging in head and neck cancer, Eur. Radiol. 20 (2010) 2358–2366, https://doi.org/10.1007/s00330-010-1803-x.

[28] S. Braun, F.D. Vogl, B. Naume, W. Janni, M.P. Osborne, R.C. Coombes, G. Schlimok, I.J. Diel, B. Gerber, G. Gebauer, J.-Y. Pierga, C. Marth, D. Oruzio, G. Wiedswang, E.-F. Solomayer, G. Kundt, B. Strobl, T. Fehm, G.Y.C. Wong, J. Bliss, A. Vincent-Salomon, K. Pantel, A pooled analysis of bone marrow micrometastasis in breast cancer, N. Engl. J. Med. 353 (2005) 793–802, https://doi.org/10.1056/NEJMoa050434.

[29] K. Bartkowiak, M. Kwiatkowski, F. Buck, T.M. Gorges, L. Nilse, V. Assmann, A. Andreas, V. Müller, H. Wikman, S. Riethdorf, H. Schlüter, K. Pantel, Disseminated tumor cells persist in the bone marrow of breast cancer patients through sustained activation of the unfolded protein response, Cancer Res. 75 (2015) 5367–5377, https://doi.org/10.1158/0008-5472.CAN-14-3728.

[30] S. Riethdorf, K. Pantel, Disseminated tumor cells in bone marrow and circulating tumor cells in blood of breast cancer patients: current state of detection and characterization, Pathobiology 75 (2008) 140–148, https://doi.org/10.1159/000123852.

[31] S.G. Pillai, S. Li, C.M. Siddappa, M.J. Ellis, M.A. Watson, R. Aft, Identifying biomarkers of breast cancer micrometastatic disease in bone marrow using a patient-derived xenograft mouse model, Breast Cancer Res. 20 (2018) 2, , https://doi.org/10.1186/s13058-017-0927-1.

[32] V. Ribrag, D. Vanel, S. Leboulleux, J. Lumbroso, D. Couanet, G. Bonniaud, A. Aupérin, F. Masson, J. Bosq, V. Edeline, C. Fermé, F. Pigneur, M. Schlumberger, Prospective study of bone marrow infiltration in aggressive lymphoma by three independent methods: whole-body MRI, PET/CT and bone marrow biopsy, Eur. J. Radiol. 66 (2008) 325–331, https://doi.org/10.1016/j.ejrad.2007.06.014.

[33] G. Cheng, W. Chen, W. Chamroonrat, D.A. Torigian, H. Zhuang, A. Alavi, Biopsy versus FDG PET/CT in the initial evaluation of bone marrow involvement in pediatric lymphoma patients, Eur. J. Nucl. Med. Mol. Imaging 38 (2011) 1469–1476, https://doi.org/10.1007/s00259-011-1815-z.

[34] J. Kotzerke, A. Guhlmann, F. Moog, N. Frickhofen, S.N. Reske, Role of attenuation correction for fluorine-18 fluorodeoxyglucose positron emission tomography in the primary staging of malignant lymphoma, Eur. J. Nucl. Med. 26 (1999) 31–38.

[35] K.M. Al-Muqbel, Bone marrow metastasis is an early stage of bone metastasis in breast Cancer detected clinically by F18-FDG-PET/CT imaging, Biomed. Res. Int. 2017 (2017) 9852632, , https://doi.org/10.1155/2017/9852632.

[36] L. Evangelista, A. Panunzio, R. Polverosi, A. Ferretti, S. Chondrogiannis, F. Pomerri, D. Rubello, P.C. Muzzio, Early bone marrow metastasis detection: the additional value of FDG-PET/CT vs. CT imaging, Biomed. Pharmacother. 66 (2012) 448–453, https://doi.org/10.1016/j.biopha.2012.06.004.

[37] G.C. Cawley, N.L.C. Talbot, Fast exact leave-one-out cross-validation of sparse least-squares support vector machines, Neural Netw. 17 (2004) 1467–1475, https://doi.org/10.1016/j.neunet.2004.07.002.

[38] W.H. Hampton, N. Asadi, I.R. Olson, Good things for those who wait: predictive modeling highlights importance of delay discounting for income attainment, Front. Psychol. 9 (2018) 1545, , https://doi.org/10.3389/fpsyg.2018.01545.

[39] M.N.S. Zainudin, M.N. Sulaiman, N. Mustapha, T. Perumal, R. Mohamed, Recognizing Complex Human Activities Using Hybrid Feature Selections Based on an Accelerometer Sensor, (2017).

[40] H.-L. Chen, G. Wang, C. Ma, Z.-N. Cai, W.-B. Liu, S.-J. Wang, An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson's disease, Neurocomputing 184 (2016) 131–144, https://doi.org/10.1016/J.NEUCOM.2015.07.138.

[41] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer New York, New York, NY, 2009, https://doi.org/10.1007/978-0-387-84858-7.

[42] L. Wang, Y. Wang, Q. Chang, Feature selection methods for big data bioinformatics: a survey from the search perspective, Methods 111 (2016) 21–31, https://doi.org/10.1016/j.ymeth.2016.08.014.