

UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Klinik und Poliklinik für
Neuroradiologische Diagnostik und Intervention

Direktor: Prof. Dr. med. Jens Fiehler

**Discordant and Converting Receptor Expressions in Brain Metastases from
Breast Cancer: MRI-Based Non-Invasive Receptor Status Tracking**

Dissertation

zur Erlangung des Grades eines Doktors der Medizin
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Alexander Joachim Giesbert Heitkamp

aus Werne

Hamburg 2023

**Angenommen von der
Medizinischen Fakultät der Universität Hamburg am: 12.04.2024**

**Veröffentlicht mit Genehmigung der
Medizinischen Fakultät der Universität Hamburg.**

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Katja Weisel

Prüfungsausschuss, zweite/r Gutachter/in: PD Dr. Susanne Gellißen

Inhaltsverzeichnis

1.	Originalveröffentlichung	4
2.	Einleitung.....	21
2.1	Radiomics und Machine Learning in der medizinischen Bilddatenanalyse	21
2.2	Hirnmetastasen von Brustkrebs: Klinische Implikationen	23
3.	Material und Methoden	24
3.1	Patientenkollektiv	24
3.2	Bildgebung, Segmentierung, Bildpostprozessierung	26
3.3	Extraktion der quantitativen Bildmerkmale und Machine Learning	27
3.4	Selektion quantitativer Bildmerkmale, Modellhyperparameter	27
3.5	Erstellung der Trainings- und Testsets	28
3.6	Statistik	28
4.	Ergebnisse.....	29
4.1	Studienkollektiv	29
4.2	Östrogenrezeptor	29
4.3	Progesteronrezeptor	29
4.4	Humaner EGF-Rezeptor	29
4.5	Merkmalswichtigkeit und radiomische Signatur	30
5.	Diskussion	31
6.	Schlussfolgerung	35
7.	Zusammenfassung	36
8.	Abstract.....	37
9.	Abkürzungsverzeichnis	38
10.	Literaturverzeichnis	39
11.	Erklärung des Eigenanteils	44
12.	Danksagung	45
13.	Lebenslauf	46
14.	Eidesstattliche Erklärung	47

Article

Discordant and Converting Receptor Expressions in Brain Metastases from Breast Cancer: MRI-Based Non-Invasive Receptor Status Tracking

Alexander Heitkamp ^{1,*}, Frederic Madesta ², Sophia Amberg ¹, Schohla Wahaj ³, Tanja Schröder ¹, Matthias Bechstein ¹, Lukas Meyer ¹, Gabriel Broocks ¹, Uta Hanning ¹, Tobias Gauer ³, René Werner ^{2,4}, Jens Fiehler ¹, Susanne Gellißen ¹ and Helge C. Kniep ¹

¹ Department of Diagnostic and Interventional Neuroradiology, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

² Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

³ Department of Radiotherapy and Radiation Oncology, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

⁴ Center for Biomedical Artificial Intelligence (bAIome), University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany

* Correspondence: alexander.heitkamp@stud.uke.uni-hamburg.de; Tel.: +49-(0)-151-511-37933

Simple Summary: Discordance and conversion of receptor expressions in metastatic lesions and primary tumors is often observed in patients with brain metastases from breast cancer. Personalized therapy requires continuous monitoring of receptor expressions and dynamic adaptation of applied targeted treatment options. This study sought to evaluate if quantitative MR image features can predict the receptor status of brain metastases from breast cancer using machine learning algorithms. Results indicate that receptor status can be differentiated non-invasively based on routine MR imaging data. The proposed approach could allow non-invasive expression tracking at high frequencies, and may support dynamic treatment optimization for personalized therapies.

Abstract: Discordance and conversion of receptor expressions in metastatic lesions and primary tumors is often observed in patients with brain metastases from breast cancer. Therefore, personalized therapy requires continuous monitoring of receptor expressions and dynamic adaptation of applied targeted treatment options. Radiological in vivo techniques may allow receptor status tracking at high frequencies at low risk and cost. The present study aims to investigate the potential of receptor status prediction through machine-learning-based analysis of radiomic MR image features. The analysis is based on 412 brain metastases samples from 106 patients acquired between 09/2007 and 09/2021. Inclusion criteria were as follows: diagnosed cerebral metastases from breast cancer; histopathology reports on progesterone (PR), estrogen (ER), and human epidermal growth factor 2 (HER2) receptor status; and availability of MR imaging data. In total, 3367 quantitative features of T1 contrast-enhanced, T1 non-enhanced, and FLAIR images and corresponding patient age were evaluated utilizing random forest algorithms. Feature importance was assessed using Gini impurity measures. Predictive performance was tested using 10 permuted 5-fold cross-validation sets employing the 30 most important features of each training set. Receiver operating characteristic areas under the curves of the validation sets were 0.82 (95% confidence interval [0.78; 0.85]) for ER+, 0.73 [0.69; 0.77] for PR+, and 0.74 [0.70; 0.78] for HER2+. Observations indicate that MR image features employed in a machine learning classifier could provide high discriminatory accuracy in predicting the receptor status of brain metastases from breast cancer.

Keywords: metastatic cancer; brain metastases; receptor status; machine learning; radiomics



Citation: Heitkamp, A.; Madesta, F.; Amberg, S.; Wahaj, S.; Schröder, T.; Bechstein, M.; Meyer, L.; Broocks, G.; Hanning, U.; Gauer, T.; et al. Discordant and Converting Receptor Expressions in Brain Metastases from Breast Cancer: MRI-Based Non-Invasive Receptor Status Tracking. *Cancers* **2023**, *15*, 2880. <https://doi.org/10.3390/cancers15112880>

Academic Editor:
Ulrich Ronellenfisch

Received: 4 April 2023
Revised: 15 May 2023
Accepted: 18 May 2023
Published: 23 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The occurrence of brain metastases substantially impacts prognosis and health-related quality of life in patients with breast cancer [1]. For breast cancer, reported incidences for cerebral metastases range between 10–30% [2–4]. In breast cancer therapy, estrogen receptor alpha (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status of the primary tumor have become the major determinants of targeted treatment regimens [5]. Moreover, in the case of metastatic disease, therapeutic decisions are often based on the receptor expression of the primary lesion. Biopsies are seldom performed, and are only considered if the metastases' locations and the patient's general health condition allow invasive biopsy procedures. Several studies have demonstrated that metastases have a substantial tendency for mutations in receptor expression, leading to discordant and diverging receptor status [6–14]. A recent meta-analysis reported the following rates for receptor conversion in patients with metastatic breast cancer by primary tumor receptor expression: For ER, PR, and HER2, the pooled positive to negative conversion rates were 22.5%, 49.4%, and 21.3%, respectively; negative to positive conversion rates were 21.5%, 15.9%, and 9.5% [7]. Following these findings, several guidelines included recommendations for biopsies of metastases to confirm the metastases' receptor expressions [5]. Moreover, recent updates extend this approach, and also recommend reassessments of metastatic receptor expressions if clinically feasible [1]. However, for cerebral metastases, biopsies are associated with risks of hemorrhage, infection, and neurologic complications, or are technically not feasible [7].

For these patients, non-invasive advanced radiological techniques may allow receptor status tracking at high sampling rates at low risk and cost. Such an approach could enable continuous monitoring of receptor expressions and dynamic adaptation of applied targeted treatment options. The present study investigates the potential of receptor status determination via a machine learning approach. We hypothesize that quantitative imaging biomarkers of brain metastases extracted from standard-of-care multispectral MR image data can be used to non-invasively determine the receptor status of these metastatic lesions.

2. Materials and Methods

This single-center retrospective study based on anonymized data was approved by the local ethics committee and the requirement for patient consent was waived (WF-144/20). The study was conducted in accordance with the Declaration of Helsinki. The data supporting the results of the analysis are available from the corresponding author upon reasonable request.

An overview of the proposed algorithm for receptor status determination is given in Figure 1; its modules are described in the following.

2.1. Patients

MR imaging studies over a 14-year period (September 2007–September 2021) were retrospectively screened for the presence of brain metastases based on availability sampling. Clinical parameters were collected through chart review. Inclusion criteria (Figure 2) were defined as follows: (a) intra-axial tumors; (b) MR image acquisition; (c) existing pathology report; (d) age > 18 years. Subsequently, the patients were further screened. Exclusion criteria (Figure 2) were defined as follows: (a) primary brain tumor; (b) >1 primary tumor type; (c) other tumor type than breast cancer; (d) not all MRI sequences available; (e) spatial resolution too low; (f) previous metastasis-specific therapy; (g) artifacts; (h) not every receptor status available. Included images were acquired during the initial diagnostic procedure and before specific therapies for the brain metastases were initiated. In total, 412 brain metastases from 106 patients (mean age 56 year, range 33–84 year; 103 females, 3 males) were included (Table 1). Extracted patient data included type of primary neoplasm, receptor status, patient sex, and patient age. In order to avoid a clumping effect with regard to the feature expression and machine learning model training, the number of metastases per patient was limited to the 10 largest lesions. We obtained the exact tumor classification

from the pathology department of our hospital from the tissue biopsies of the primary tumor side or of the metastases. For the classification of receptor status, our laboratory works according to the AWMF guidelines. Thus, a negative ER/PR receptor status occurs when there are less than 1% positive tumor cells, a low positive ER/PR status occurs when there are 1–9% positive tumor cells, and a positive status occurs when there are 10% or more than 10% positive tumor cells in the tissue examined [15,16]. Regarding our study, we considered a low positive status as the threshold for ER+/PR+. HER2 positivity is indicated by a uniformly intense circular membrane response of more than 10% (=3+) of tumor cells during immunohistochemistry [17], or a preferably detected gene amplification by in situ hybridization (ISH). Of the 106 patients included, metastasis-specific receptor status was known in 38 patients. In the remaining collective, the receptor status of the primary tumor was used.

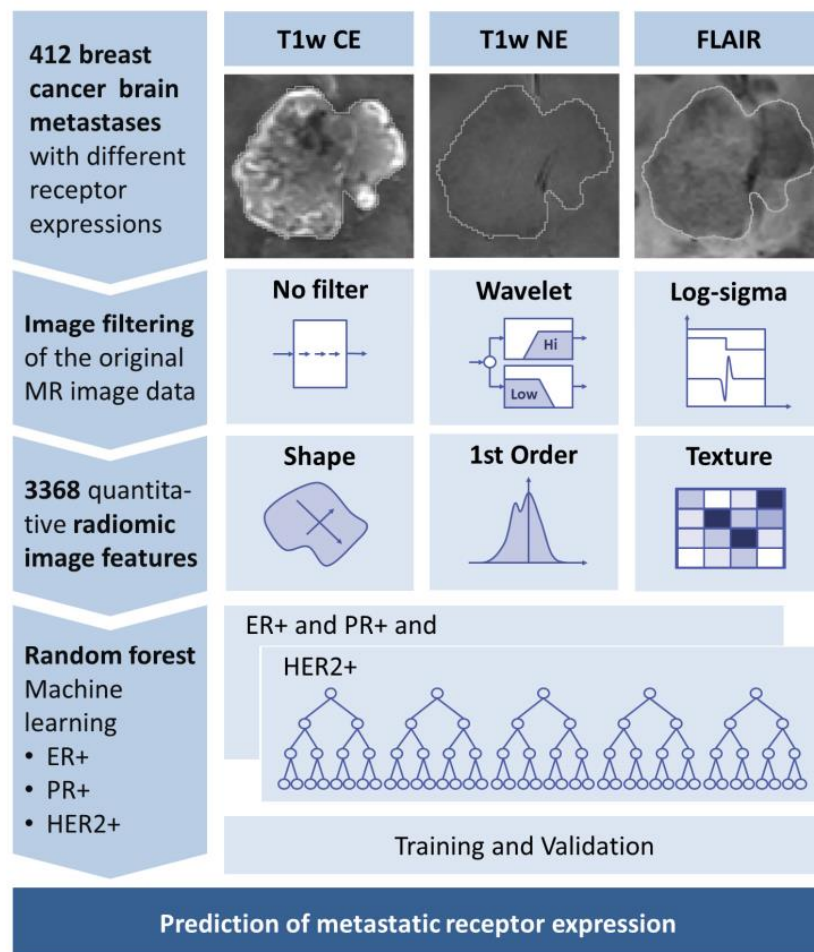


Figure 1. Overview of proposed algorithm for receptor status prediction. T1w CE, contrast-enhanced T1-weighted MR images; T1w NE, non-contrast-enhanced T1-weighted MR images; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2.

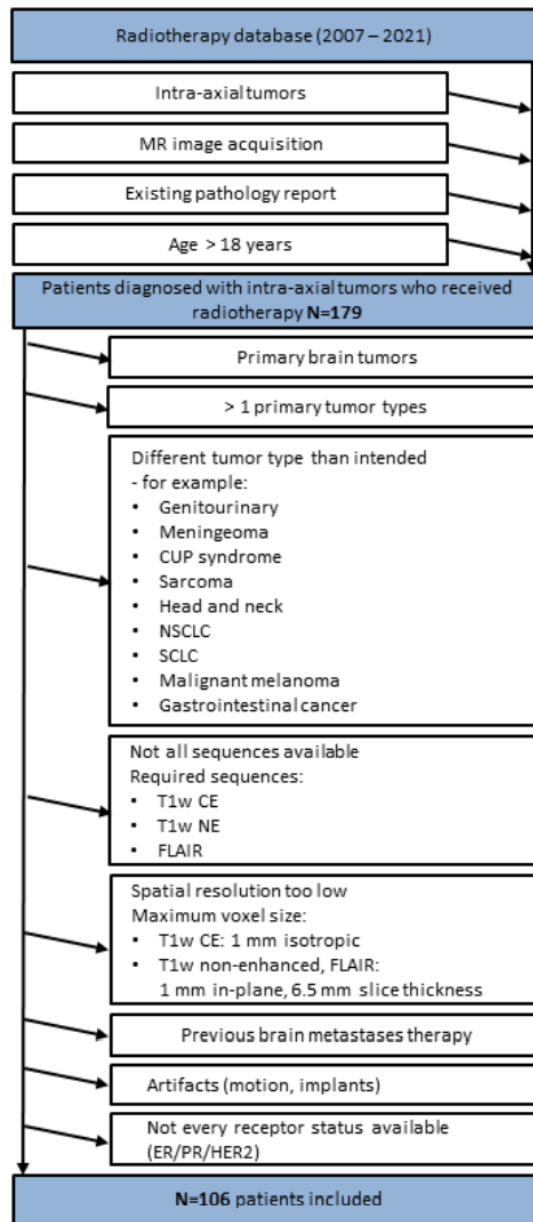


Figure 2. Patient inclusion and exclusion criteria flow chart. T1w, T1-weighted; CE, contrast-enhanced; NE, non-contrast-enhanced; FLAIR, non-contrast axial-fluid-attenuated inversion recovery; CUP, cancer of unknown primary; NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2.

Table 1. Demographic and clinical characteristics of the included metastases samples. The table shows number of included metastases, number of respective patients, age at the time of brain metastases diagnosis by receptor status, and number of metastases of the sample patient by receptor status. ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2. * Number of patients included with single brain metastasis; ** number of patients included with multiple brain metastases.

	Number of		Age of Sample Patient			Number of Metastases of Sample Patient		
	Metastases	Patients	Mean	Min	Max	Mean	Min	Max
All patients	412	106 30 * 76 **	55.9	33	84	5.4	1	66
Classifier 1 (ER)	Total samples							
ER+	190	53	58.3	33	76	4	1	63
ER−	222	53	53.9	36	79	5.9	1	66
Classifier 2 (PR)	Total samples							
PR+	147	37	59.5	39	76	4.5	1	63
PR−	265	69	53.9	33	84	5.7	1	66
Classifier 3 (HER)	Total samples							
HER+	169	45	54.5	36	81	4.1	1	63
HER−	243	61	56.9	33	84	6.1	1	66

2.2. Image Acquisition

Images were acquired with 1.5 Tesla scanners (Magnetom® Symphony, Magnetom® Avanto, and Magnetom® Sonata, all manufactured by Siemens Healthcare, Erlangen, Germany) or 3 Tesla scanners (Ingenia®, Philips Medical Systems, Best, The Netherlands; Magnetom® Skyra and Magnetom® Vida, Siemens Healthineers AG, Erlangen, Germany). Since some MRI studies were acquired in external clinics, the scanners could not be traced for all included patients. Employed imaging protocols included non-enhanced axial T1-weighted (T1w NE) spin echo with flow compensation and non-contrast axial-fluid-attenuated inversion recovery (FLAIR). After weight-adjusted injection of a gadolinium-based contrast agent (0.1 mmol/kg), contrast-enhanced axial T1-weighted spin echo sequences with flow compensation or three-dimensional T1w gradient echo sequences were acquired (T1w CE).

2.3. Tumor Segmentation

Tumor delineation was performed based on the original T1w CE scans using a semi-automatic approach facilitated by ITK-SNAP software (Biomedical Imaging Resource; version number 3.8.0; Penn Image Computing and Science Laboratory University of Pennsylvania (PA), Scientific Computing and Imaging Institute University of Utah (UT)) and controlled by two radiologists (SG, 13 years clinical experience in diagnostic neuroradiology in an academic full-service hospital; HK, 2 years clinical experience in diagnostic neuroradiology in an academic full-service hospital). Regions of interest (ROIs) were defined according to the visible gross tumor volume; readers were blinded to all clinical data. Consensus ROIs were calculated based on overlapping segmentations when inconsistencies occurred.

2.4. Image Post-Processing

In order to extract ROIs from all acquired sequences and to normalize texture metrics, the bone mass was removed from all images with ROBEX software (NITRC; University

of Massachusetts Medical School in Worcester, MA) [18], and all images were registered to a common standard space (MNI152 [19]). To further reduce bias in texture analysis, spatial resolutions of target spaces for image registration were selected so that no upsampling of image data was necessary. Hence, images were registered to atlases resliced to the maximum voxel dimensions of each spatial dimension in the respective unprocessed original data. Registration was conducted with NiftyReg (Centre for Medical Image Computing; University College London, UK) [20]. To account for non-standardized intensity calculations of MR scanners, all images were N4ITK bias-corrected [21] and piecewise linear histogram matched [22]. Tumor segmentations were transformed into MNI standard space employing transformation matrices obtained from image registration. Two MDs (HK and SG) visually verified the success of image post-processing. An overview of the image post-processing steps is given in Figure 3.

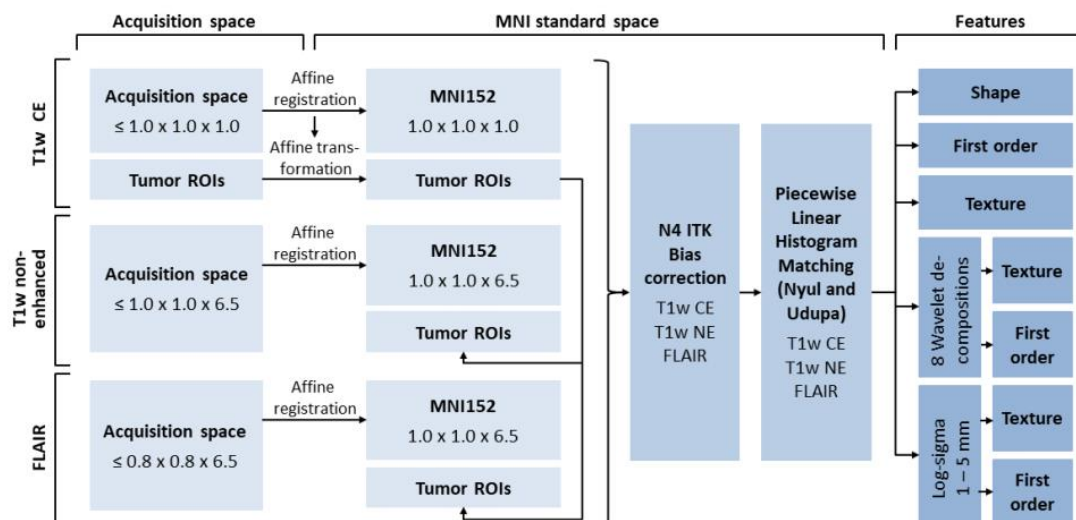


Figure 3. Image processing pipeline. Spatial resolutions are shown in mm. MNI, Montreal Neurological Institute of McGill University Health Centre; T1w, T1-weighted; CE, contrast-enhanced; NE, non-enhanced; FLAIR, non-contrast axial-fluid-attenuated inversion recovery; ROI, region of interest.

2.5. Feature Extraction

Texture- and intensity-based analyses require a certain minimum number of included voxels to ensure meaningful marker results. Hence, only features from metastatic lesions fulfilling the following criteria were included in the analysis: (a) metastasis volume equal to or greater than 100 mm^3 , equating to a cube with an edge length of 4.6 mm; (b) metastasis belongs to the 10 largest metastases per patient. The total number of included metastases was $n = 412$ (Table 1). Extracted radiomic features were defined according to the employed PyRadiomics Python package v2.2.0 [23]. For each analyzed MR sequence (T1w CE, FLAIR, T1w non-enhanced), extracted features comprised 252 first-order features (18 based on the unfiltered images, 144 wavelet decompositions, 90 log-sigma Laplacian of Gaussian-filtered images), 952 texture features (68 based on the unfiltered images, 544 wavelet decompositions, 340 log-sigma Laplacian of Gaussian-filtered images), and 14 shape features (based on the unfiltered images). Since some features in some sequences could not be calculated from the image properties, these features were excluded from further consideration, resulting in a total number of 3368 image features and the patient age.

2.6. Machine Learning

Machine-learning-based classification of receptor status was performed using random forest algorithms implemented in the Python scikit-learn environment v1.1.1 [24]. Random

forest classifiers have a comparably low tendency to overfit [25,26], and support classification tasks for data sets comprising numerous and heterogeneous predictors. Furthermore, random forest classifiers allow analysis of cluster-correlated observations, e.g., patients with multiple metastases [27]. Three different classifiers were trained for differentiation of the following cases: Classifier 1, hormone/estrogen-receptor-positive vs. -negative; Classifier 2, hormone/progesterone-receptor-positive vs. -negative; Classifier 3, human epidermal growth factor receptor 2 (HER2)-positive vs. -negative. Triple-negative status can be assumed for ER-, PR-, and HER2-negative metastases.

2.7. Feature Selection

Predictive values of the employed image features were calculated separately for each training data set based on Gini impurity measures [28]. Since many features are expected to be correlated, e.g., due to using several filters on the same image, and owing to the risk of overfitting the model with too many features, we restricted the number of features used to the 30 most important features [29]. For all model evaluation runs, feature importance was determined based on the respective training data only.

2.8. Model Hyperparameters, Training, and Testing

The number of trees was set to random forest classifier default. The square root of the total number of features utilized for training and prediction was used for setting the number of features per node according to the literature [25,30]. In total, 10 randomly permuted 5-fold cross-validation sets [31,32] generated from the $n = 412$ samples were used for model evaluation. All cross-validation sets were randomly drawn with stratification for receptor status. In addition, to prevent bias from cluster correlation, metastases of patients with multiple lesions were either assigned to the validation or to the training set. Predictive performance was evaluated based on all validation/test sets, each set containing $n = 329$ training samples and $n = 83$ validation/test samples.

2.9. Statistics

Receiver operating characteristic (ROC) curves of the training/validation sets were calculated based on cross-validation set means. Hence, ROC curves from cross-validation may serve as unbiased estimates for model classification performance in a generalized setting. Statistical significance of the areas under the curves (AUCs) was assumed if p -values were < 0.05 in all cross-validation sets. Model prediction instability (i.e., standard deviation of ROC curves) was evaluated using 10 randomly drawn 5-fold cross-validation sets. p -values were calculated with Mann–Whitney/Wilcoxon U statistics using the verification v1.42 R package [33]. Classifier performance was evaluated through ROC AUC metrics. Furthermore, the classifiers were analyzed at operating points according to the maximum value of the Matthews correlation coefficient (MCC) [34]. The MCC evaluates all fields of the confusion matrix, and is considered as a favorable measure for unbiased comparisons of binary classifiers [35]. Incorporating TP s (true positives), TN s (true negatives), FP s (false positives), and FN s (false negatives), MCC is defined as

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Confidence intervals for classification performance metrics were calculated using pROC v1.18.0 [36], ThresholdROC v2.9.0, and psychometric v2.3 R packages, and the Python scikit-learn metrics environment v1.1.1 [24].

3. Results

The analysis is based on T1w CF, FLAIR, and T1w NF MR images of 412 brain metastases from breast cancer. The training set includes $n = 329$ samples; the independent validation set includes $n = 83$ samples in 1 out of 10 randomly drawn 5-fold cross-validation sets.

3.1. ER Receptor

ROC AUC in the validation sets for predicting the receptor status based on multiparametric MR scans was 0.82 with 95% CI [0.78; 0.85] (Figure 4). At the maximum MCC cut-off point, the classifier yielded specificities of 68% (95% CI [65%; 71%]) at sensitivities of 84% (95% CI [81%; 88%]) (Figure 4, Table 2). Furthermore, the positive predictive value (PPV) and negative predictive value (NPV) were 69% and 84%, respectively; accuracy was 76% and MCC was 0.53 (Table 2).

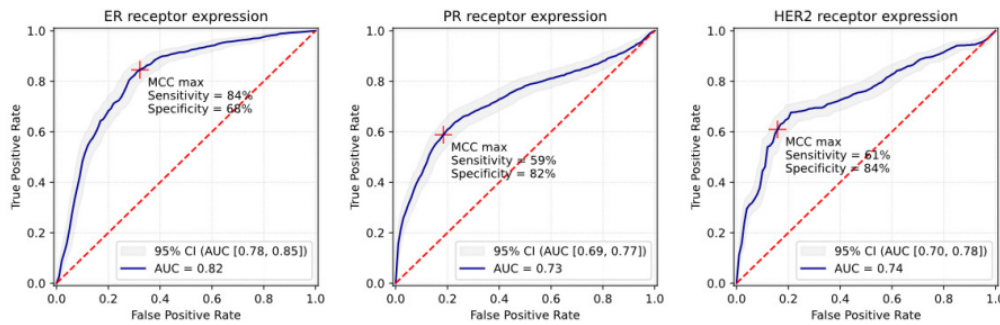


Figure 4. Receiver operating characteristic (ROC) curves and Matthews correlation coefficient (MCC) cut-off point analysis of ER+/ER+ and HER2+ classifiers. Training/validation: 5-fold cross-validation, consideration of cluster correlation, sample size of 329 metastases. Independent test/validation set: sample size of 83 metastases. Crosses depict the maximum MCC cut-off points. ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; AUC, area under the curve; CI, confidence interval.

Table 2. Classification performance metrics of independent validation sets. Classification performance metrics of the machine learning classifiers. Classifier metrics are shown at cut-off points according to the maximum Matthews correlation coefficient. AUC, area under the curve; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; PPV, positive predictive value; NPV, negative predictive value; MCC, Matthews correlation coefficient; CI, confidence interval.

Receptor Status	Matthews Correlation Coefficient (MCC) Maximum Operating Point						
	AUC [95% CI]	Sensitivity [95% CI]	Specificity [95% CI]	PPV [95% CI]	NPV [95% CI]	Accuracy [95% CI]	MCC [95% CI]
5-fold cross-validation (n = 412 samples)							
ER+	0.82 [0.78; 0.85]	84% [81%; 88%]	68% [65%; 71%]	69% [67%; 72%]	84% [82%; 88%]	76% [73%; 78%]	0.53 [0.49; 0.58]
PR+	0.73 [0.69; 0.77]	59% [56%; 64%]	82% [77%; 85%]	69% [67%; 76%]	74% [70%; 76%]	72% [69%; 75%]	0.42 [0.37; 0.48]
HER2+	0.74 [0.70; 0.78]	61% [57%; 65%]	84% [81%; 88%]	73% [72%; 80%]	76% [74%; 78%]	75% [73%; 76%]	0.47 [0.45; 0.52]

3.2. PR Receptor

ROC AUC of the validation sets was 0.73 with 95% CI [0.69; 0.77] (Figure 4). At the maximum MCC cut-off point (MCC = 0.42), specificities of 82% (95% CI [77%; 85%]) at sensitivities of 59% (95% CI [56%; 64%]) were observed (Figure 4, Table 2). PPV and NPV were 69% and 74%, respectively; accuracy was 72% (Table 2).

3.3. HER2 Receptor

ROC AUC of the validation sets was 0.74 with 95% CI [0.70; 0.78] (Figure 4). At the maximum MCC cut-off point (MCC = 0.47), specificities of 84% (95% CI [81%; 88%]) at sensitivities of 61% (95% CI [57%; 65%]) were observed (Figure 4, Table 2). PPV was 73%, NPV was 76%, and accuracy was 75% (Table 2).

3.4. Triple-Negative

Triple-negative status can be assumed for metastases with PR-, ER-, and HER2-negative predictions.

Comparative analysis of randomly drawn five-fold cross-validation sets confirmed low variability of results, hence indicating stable model performance. ROC AUCs of the independent validation/test sets were >0.5 with p -value < 0.05 for all three classifiers.

3.5. Feature Importance

Mean importance numbers generated from all training sets were used for the feature importance analysis. While PR+ and HER2+ classifiers employed mainly T1w NE images (36% of total predictive importance), ER+ classification was mostly FLAIR-dependent (45% of total predictive importance). For T1w CE (30%), a lower contribution to predictive performance was observed in PR+ and HER+ classifiers, respectively; 29% was observed for the ER+ classifier (Figure 5). Overall, however, the different image sequences have similar proportions of the feature importance distribution, especially with regard to PR+ and HER2+. Top 100 ranked features differed between ER+ and PR+/HER2+ classifiers, indicating the presence of specific predictive imaging marker signatures for ER+ classifiers and similar predictive imaging markers for PR+ and HER2+ classifiers. In total, only 2 identical image features ranked among the top 100 most important predictors for all 3 receptor classes, but considering only PR+ and HER2+ receptor classes, they share 98 of the 100 most important image features. With regards to applied filters, the shares in feature importance contribution for ER+ and PR+/HER2+ classifiers were as follows: Log-sigma-filtered images, 27% vs. 36%; wavelet-filtered images, 67% vs. 57%; and unfiltered original images, 6% vs. 7% (Figure 5). Differentiated by feature classes, the ER+ classifier utilizes more texture features (78%) compared to the PR+/HER2+ classifiers (75%). Shape features were not ranked among the top 100 predictive markers for the ER+ classifier, and had only little impact in PR+/HER2+ classification (Figure 5). Overall, the lowest-performing image sequence was T1w CE, the lowest-performing feature class was shape, and including log-sigma and wavelet filters in the analysis seems to be important (Figure 5).

Figure 6 shows boxplots of the top 10 normalized feature values for the ER+ classifier (A), PR+ classifier (B), and for the HER2+ classifier (C). The radiomic signature value distribution suggests that ER+ metastases have lower gray level emphasis compared to hormone-receptor-negative metastases; however, zone entropy is more pronounced in ER+ metastases, but both zone entropy predictors exhibit large variability for the same predictor set, and might not allow for stable classification results (Figure 6A). Figure 6B,C confirm similar predictor distributions for PR+ and HER2+ metastases, but the overall picture of the distributions of individual predictors is more heterogeneous compared to the ER metastases. Still, different characteristics of gray level emphasis seem to be important in PR+ prediction as well. The first-order mean absolute deviation appears to be much higher for PR+ and HER2+ metastases than for negative receptor status, exhibiting large variability for the same predictor set. All differences in mean feature values between receptor-positive and -negative metastases were statistically significant (p -value < 0.05).

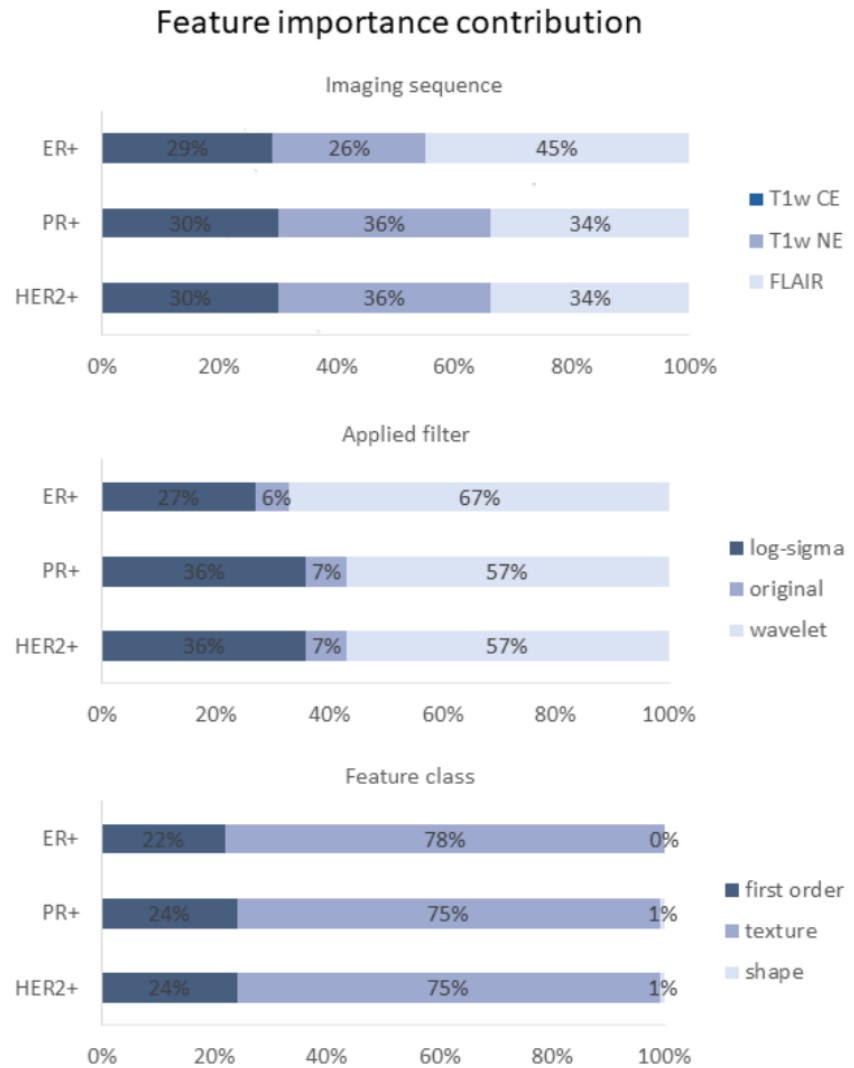


Figure 5. Radiomic feature distributions of top 100 ranked most important predictors for HER2+ and ER+/PR+ classifiers. Numbers show shares of total feature importance contribution. ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; T1w, T1-weighted; CE, contrast-enhanced; NE, non-enhanced; FLAIR, non-contrast axial-fluid-attenuated inversion recovery.

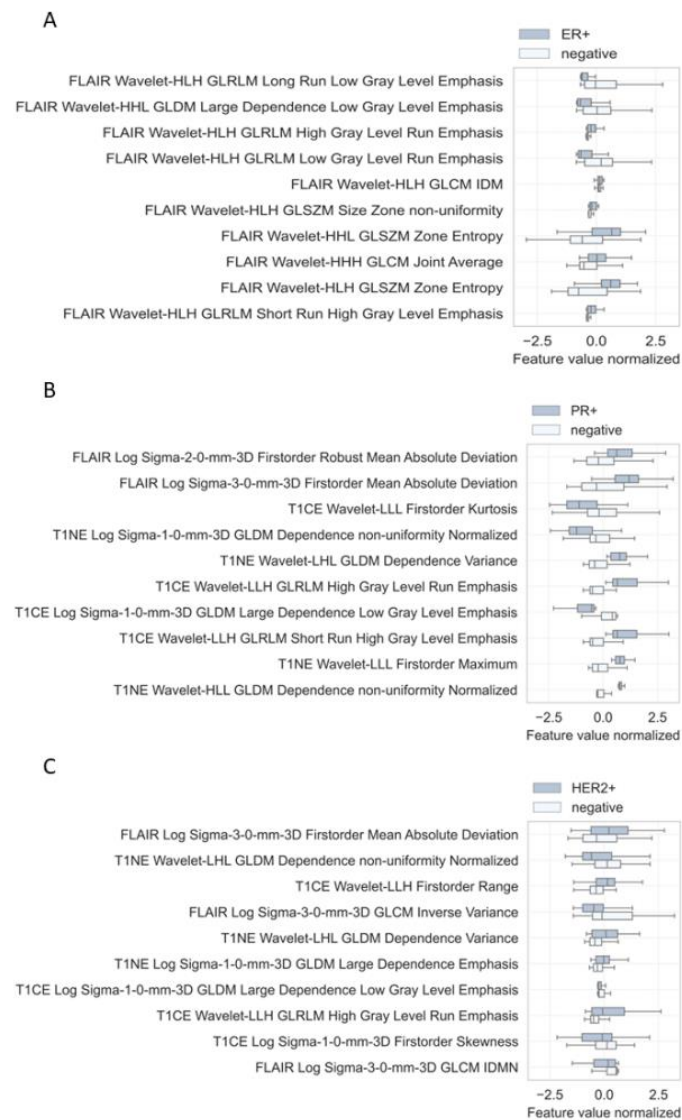


Figure 6. Radiomic signatures of 10 most important image features for the ER+ classifier (A), PR+ classifier (B), and HER2+ classifier (C). Comparison of respective receptor-positive metastases with receptor-negative metastases. T1NE, T1 non-enhanced; T1CE, T1 contrast-enhanced; FLAIR, non-contrast axial-fluid-attenuated inversion recovery; H, high-pass wavelet decomposition; L, low-pass wavelet decomposition.

4. Discussion

We proposed to use quantitative image features from standard multiparametric MR scans to determine the receptor status of brain metastases in patients with metastatic breast cancer without invasive procedures. Our results show that machine learning algorithms can predict the receptor status, achieving AUCs of 0.82 for ER+ metastases, 0.73 for PR+ metastases, and 0.74 for HER2+ metastases. The study demonstrates that an artificial intelligence approach has the potential to non-invasively track receptor expression in patients with metastatic breast cancer.

Over the past few years, advances in targeted treatment options for patients with primary and advanced breast cancer have made receptor status expression of the primary tumor a leading determinant of therapeutic regimes, and have led to a significant improvement in outcomes [1,37,38]. However, as indicated above, in patients with metastatic breast cancer, diverging and discordant receptor status are observed with the following frequencies: 19.3% for ER, 30.9% for PR, and 10.3% for HER2 expression [7]. Moreover, studies have shown that receptor expressions can diverge even for different metastases in the same patient [8]. Receptor divergence might have significant impact in both of the two possible scenarios: In the case where the receptor expression is lost, patients will receive their now-ineffective targeted therapy at the price of related adverse effects and additional costs for the medication. In the case of mutations leading to receptor expression gain, patients will not be prescribed respective medications and will not benefit from effective targeted treatment options. However, to date, there is no randomized, controlled, trial-based evidence for improved survival in patients with therapy adaptations according to their metastases' receptor status [1]. Nevertheless, in order to bridge this evidence gap and to ensure the best possible patient care, recent guidelines recommend (a) to assess and to retest receptor status of the metastatic lesion at least once [5,39], and (b) to align therapeutic regimes according to the receptor expression of the metastatic lesion [5], or to consider any positive receptor expression from metastases or primary tumor biopsies [1]. Still, biopsies and rebiopsies can only be conducted if the respective lesion is accessible and adverse effects are expected at appropriate rates considering the potential benefits of therapy adjustments. Especially in patients with brain metastases, both of these preconditions are often difficult to evaluate and are seldom fulfilled. Hence, the here-proposed method for non-invasive receptor status determination based on imaging biomarkers derived from standard-of-care MR brain scans might (a) improve therapeutic management by allowing receptor status monitoring at high frequencies and enabling agile therapy optimization, (b) enhance the general understanding of receptor divergence mechanisms, and (c) generate further evidence with respect to the effects of receptor divergence and applied targeted treatment regimens on overall survival.

Several studies have investigated approaches for non-invasive receptor status diagnosis in breast cancer. However, most analyses were limited to primary tumor characteristics: Catalano et al. (2017) [40] show correlations of breast cancer phenotype with imaging markers from PET-MR scans, with cross-validation accuracies of 38% to 48% in a multivariate analysis. Incoronato et al. (2018) [41] also utilize PET-MR imaging biomarkers, and report multivariate model accuracies of 78%; however, cross-validation was not performed, and results might worsen significantly in a training/validation approach. With ROC AUCs of 0.62 to 0.65, a large MR-imaging-based analysis by Saha et al. (2018) [42] reported moderate associations of molecular subtypes with imaging biomarkers.

To date, only a few approaches for non-invasive metastases receptor expression diagnosis have been investigated. An analysis related to the receptor status of breast cancer metastases in the brain was conducted by Luo et al. (2022) [43]: A logistic regression model was trained to predict the receptor status of brain metastases. The predictions were compared to pathology reports of the metastases and primary tumors. The study reported AUCs of 0.89, 0.88, and 0.87 for ER+, PR+, and HER2+, and receptor discordance between metastasis and primary of 27.5% for ER+/PR+ and 5% for HER2+. A training and a test set were used, but cohorts were relatively small. Another approach by Cao et al. (2022) [44] investigated whether habitat-based radiomics was capable of predicting the EGFR mutation status in brain metastases from primary lung adenocarcinoma, with an AUC of 0.9 in the best-performing external validation model. Hohm et al. (2022) [45] found that, using standardized MR imaging parameters and epidemiologic data, it is possible to detect genetic differences (H3 K27M+/-) in pediatric diffuse midline gliomas. Thus, a better prognosis may be estimated and a more personalized therapy may be offered in the future. Liquid biopsy in metastasized breast cancer as the basis for treatment decisions was discussed by Krawczyk et al. (2016) [46]. Another non-invasive approach to distinguish cerebral tumor

entities, from which receptor status prediction could also benefit, was investigated by Foda et al. (2022) [47] by examining different vascular parameters of the various tumors in MR imaging data. Nevertheless, as of today, all the approaches have not been translated into clinical routine.

The results of the training/validation analysis show superiority of the ER+ classifier (ROC AUC 0.82) compared to the PR+ and HER2+ classifiers (ROC AUC 0.73 and 0.74). Inferior results for predicting HER2 status are in line with primary tumor classification performances reported by other studies [42], suggesting the presence of more heterogeneous imaging characteristics in these lesions. A comparison of value distributions of the 10 most important histogram-based predictors employed by the ER+ classifier with value distributions of PR+ and HER2+ metastases supports this hypothesis: Respective predictors in PR+ and HER2+ lesions (Figure 6B,C) exhibit a substantial higher variability compared to ER+ tumors (Figure 6A). Furthermore, median values of positive and negative metastases are much closer together for PR+ and HER+ predictors than for ER+ metastases. This suggests inferior applicability of these high-predictive-power markers for differentiating receptor status in PR+ and HER2+ metastases, and might explain the discrepancy between selected feature sets between the three classifiers. FLAIR sequences seem to be more applicable for ER+ metastases (Figure 6A); therefore, ER+ metastases might form a specific edema-associated pattern, which can be recognized by the machine learning algorithm.

Although numerous associations between quantitative image features and molecular tissue characteristics have been demonstrated [48], radiomics-based artificial intelligence systems are still often criticized for potentially irreproducible and unintelligible decision paths. To address these concerns and establish ties to conventional image findings, we evaluated the employed image features with regards to possible interpretations in classic visual assessments. Our analysis shows that ER+ exhibit narrower value distributions and higher uniformity in wavelet-filtered images and FLAIR sequences, suggesting a more homogenous appearance of these lesions in conventional visual readings compared to hormone-receptor-negative ER+ and PR+/HER2+ metastases. Feature analysis for PR+/HER2+ metastases indicate higher image heterogeneity and more pronounced intensity edges. Top-performing predictors were similar, but often in the opposite expression direction. Triple-negative tumors were not studied separately, and due to the low overlap of shared top 100 predictors, it is difficult to draw conclusions about their radiomic signature in this study design. In line with our results, studies investigating MR imaging characteristics of primary breast tumors reported higher heterogeneity for triple-negative and HER2+ lesions compared to hormone-receptor-positive tumors [49,50]. Still, several of the observed characteristics diverge between the different analyses, and do not provide clear evidence.

Our study had general limitations of radiomic analyses, as previously reported [51–54]. Furthermore, the following study-specific limitations were observed: First, an expansion of the number of included patients would further increase the generalizability of the results. Small sample sizes are a general concern for radiomics analysis, and are due to the limited availability of standardized and annotated multicenter imaging data sets. Nevertheless, the conducted model stability assessment indicates sufficient robustness of results for the evaluation of the feasibility and limitations of imaging-marker-based receptor status predictions. Second, it cannot be excluded that some patients in this collective were already under systemic therapy during the development of the brain metastases. In these cases, therapy was not metastasis-specific like radiosurgery that would hypothetically lead to an alteration of radiomic signatures Third, we did not differentiate receptor expression by the 2015 St Gallen Consensus subtypes Luminal A, Luminal B, HER2 overexpression, and Basal-like [38,55]. Such classification would necessitate the inclusion of Ki67 prediction and a significantly more complex classifier design to account for the required multiclass approach. As the number of classes and available training sample sizes are major determinants of prediction performance, such an approach would either necessitate a substantial expansion

of sample size or would potentially degrade classifier performance. Fourth, we defined ground truth for classifier training according to pathology reports of brain metastases if available, otherwise we used primary tumor biopsies. This was due to the fact that metastases-specific pathology reports were only available for 38 of the 106 included patients. In our subgroup analysis of patients for whom we had the receptor status of primary and metastasis, a change in at least 1 receptor status (ER, PR, HER2) occurred in 8 of 38 patients, corresponding to a conversion rate of 21%. This rate is also in line with the current literature [7]. Assuming a total receptor conversion rate of 20%, respective wrongly defined ground truth in 20% of cases might partly explain the observed classification errors of 16% (ER+), 26% (PR+), and 24% (HER2+). Because the receptor status of each individual metastasis was often not determined in patients with multiple brain metastases, we used the status of the known metastasis as the ground truth in these cases. However, the notion that the status between individual metastases in the same patient is identical is not necessarily the case [8]. Nevertheless, in the subgroup of our patient collective ($n = 5$) in which the expressions of all metastases were known, they matched within patients. Fifth, low-resolution images were excluded from the study to improve the validity of the texture analysis. Inclusion of low-resolution images would potentially reduce prediction performance, as texture features substantially contribute to the model's prediction power. Sixth, the conducted semi-automatic delineation of lesion ROIs still leaves a certain degree of human observer dependence within the automated process. Consensus ROIs were partly used to minimize its impact. Further, it was observed that variations in segmentations have a comparably small effect on quantitative image feature values [56,57].

Upcoming studies may investigate the full potential of artificial-intelligence-based receptor status prediction through a substantial enlargement of the study population, employment of standardized high-resolution images, comprehensive integration of only metastases-related histopathological reports, and the adaptation of classifier design for the differentiation of subtypes according to the St Gallen Consensus. Moreover, future studies may explore other fields of application, including receptor status predictions for liver and bone metastases and adaptation of the algorithm to other primary tumors. Once approved and translated into clinical routine, non-invasive receptor status determination could significantly contribute to optimized therapeutic regimes, and may improve outcomes for patients with metastatic cancer.

5. Conclusions

We hypothesized that quantitative image features extracted from standard-of-care multiparametric MR scans can be used for non-invasive determination of the receptor status of brain metastases in patients with metastatic breast cancer. Our results suggest that machine-learning-based algorithms can predict the receptor status, with AUCs of 0.82 for ER+ metastases, 0.73 for PR+ metastases, and 0.74 for HER2+ metastases. The observed narrow confidence intervals and low standard deviations of the ROC curves indicate high stability of predictive performance in the training/validation sets. Promising AUC metrics significantly > 0.5 in the independent test set confirm the initial hypothesis. Our study demonstrates that the proposed approach could allow non-invasive receptor expression tracking in patients with metastatic breast cancer.

Future studies could explore the full potential of artificial intelligence-based receptor status prediction in brain metastases of breast cancer patients and extend this approach to other fields of interest like liver or bone metastases and all together may improve outcomes for patients with metastatic cancer.

Author Contributions: Conceptualization, L.M., U.H., J.F., S.G. and H.C.K.; Methodology, F.M., S.W., T.G. and R.W.; Software, A.H., F.M., R.W., S.G. and H.C.K.; Validation, M.B. and G.B.; Formal analysis, A.H. and T.S.; Investigation, S.A.; Resources, S.A., T.G. and R.W.; Data curation, A.H., S.W., T.S., M.B. and T.G.; Writing—original draft, A.H. and H.C.K.; Writing—review & editing, F.M., S.A., S.W., T.S., M.B., L.M., G.B., U.H., T.G., R.W., J.F. and S.G.; Visualization, A.H., F.M. and H.C.K.; Supervision,

M.B., L.M., G.B., U.H. and J.F.; Project administration, S.G. and H.C.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Approval was obtained from the ethics committee. Further requirement for consultation was waived because the study evaluates anonymized data based on retrospective analysis (WF-144/20).

Informed Consent Statement: Requirement for further consultation of the ethics committee (including requirement for informed consent) was waived by the ethics committee in the retrospective analysis of anonymized data (WF-144/20).

Data Availability Statement: The data supporting the results of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no potential conflict of interest.

References

- Cardoso, F.; Paluch-Shimon, S.; Senkus, E.; Curigliano, G.; Aapro, M.S.; Andre, F.; Barrios, C.H.; Bergh, J.; Bhattacharyya, G.S.; Biganzoli, L.; et al. 5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5). *Ann. Oncol.* **2020**, *31*, 1623–1649. [[CrossRef](#)] [[PubMed](#)]
- Wang, Y.; Ye, F.; Liang, Y.; Yang, Q. Breast cancer brain metastasis: Insight into molecular mechanisms and therapeutic strategies. *Br. J. Cancer* **2021**, *125*, 1056–1067. [[CrossRef](#)] [[PubMed](#)]
- Chamberlain, M.C.; Baik, C.S.; Gadi, V.K.; Bhatia, S.; Chow, L.Q. Systemic therapy of brain metastases: Non-small cell lung cancer, breast cancer, and melanoma. *Neuro Oncol.* **2017**, *19*, i1–i24. [[CrossRef](#)] [[PubMed](#)]
- Hosonaga, M.; Saya, H.; Arima, Y. Molecular and cellular mechanisms underlying brain metastasis of breast cancer. *Cancer Metastasis Rev.* **2020**, *39*, 711–720. [[CrossRef](#)]
- Gennari, A.; Andre, F.; Barrios, C.H.; Cortes, J.; de Azambuja, E.; DeMichele, A.; Dent, R.; Fenlon, D.; Gligorov, J.; Hurvitz, S.A.; et al. ESMO Clinical Practice Guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer. *Ann. Oncol.* **2021**, *32*, 1475–1495. [[CrossRef](#)]
- Schrijver, W.; Selenica, P.; Lee, J.Y.; Ng, C.K.Y.; Burke, K.A.; Piscuoglo, S.; Berman, S.H.; Reis-Filho, J.S.; Weigelt, B.; van Diest, P.J.; et al. Mutation Profiling of Key Cancer Genes in Primary Breast Cancers and Their Distant Metastases. *Cancer Res.* **2018**, *78*, 3112–3121. [[CrossRef](#)]
- Schrijver, W.; Suijkerbuijk, K.P.M.; van Gils, C.H.; van der Wall, E.; Moelans, C.B.; van Diest, P.J. Receptor Conversion in Distant Breast Cancer Metastases: A Systematic Review and Meta-analysis. *J. Natl. Cancer Inst.* **2018**, *110*, 568–580. [[CrossRef](#)]
- Hoefnagel, L.D.; van der Groep, P.; van de Vijver, M.J.; Boers, J.E.; Wesseling, P.; Wesseling, J.; Dutch Distant Breast Cancer Metastases, C.; van der Wall, E.; van Diest, P.J. Discordance in ERalpha, PR and HER2 receptor status across different distant breast cancer metastases within the same patient. *Ann. Oncol.* **2013**, *24*, 3017–3023. [[CrossRef](#)]
- Kroigard, A.B.; Larsen, M.J.; Thomassen, M.; Kruse, T.A. Molecular Concordance Between Primary Breast Cancer and Matched Metastases. *Breast J.* **2016**, *22*, 420–430. [[CrossRef](#)]
- Michel, A.; Oppong, M.D.; Rauschenbach, L.; Pierscianek, D.; Dinger, T.F.; Schmidt, T.; Hense, J.; Pottgen, C.; Kimmig, R.; Ahmadipour, Y.; et al. HER2 Receptor Conversion Is a strong Survival Predictor in Patients with Breast Cancer Brain Metastases. *World Neurosurg.* **2021**, *152*, e332–e343. [[CrossRef](#)]
- Yi, Z.B.; Yu, P.; Zhang, S.; Wang, W.N.; Han, Y.Q.; Ouyang, Q.C.; Yan, M.; Wang, X.J.; Hu, X.C.; Jiang, Z.F.; et al. Profile and outcome of receptor conversion in breast cancer metastases: A nation-wide multicenter epidemiological study. *Int. J. Cancer* **2021**, *148*, 692–701. [[CrossRef](#)] [[PubMed](#)]
- Aurilio, G.; Monfardini, L.; Rizzo, S.; Sciandivasci, A.; Preda, L.; Bagnardi, V.; Disalvatore, D.; Pruneri, G.; Munzone, E.; Della Vigna, P.; et al. Discordant hormone receptor and human epidermal growth factor receptor 2 status in bone metastases compared to primary breast cancer. *Acta Oncol.* **2013**, *52*, 1649–1656. [[CrossRef](#)]
- Zhao, W.; Sun, L.; Dong, G.; Wang, X.; Jia, Y.; Tong, Z. Receptor conversion impacts outcomes of different molecular subtypes of primary breast cancer. *Ther. Adv. Med. Oncol.* **2021**, *13*, 17588359211012982. [[CrossRef](#)]
- Yang, M.; Forbes, M.E.; Bitting, R.L.; O'Neill, S.S.; Chou, P.C.; Topaloglu, U.; Miller, L.D.; Hawkins, G.A.; Grant, S.C.; DeYoung, B.R.; et al. Incorporating blood-based liquid biopsy information into cancer staging: Time for a TNMB system? *Ann. Oncol.* **2018**, *29*, 311–323. [[CrossRef](#)] [[PubMed](#)]
- Nofech-Mozes, S.; Vella, E.T.; Dhesy-Thind, S.; Hagerty, K.L.; Mangu, P.B.; Temin, S.; Hanna, W.M. Systematic review on hormone receptor testing in breast cancer. *Appl. Immunohistochem. Mol. Morphol.* **2012**, *20*, 214–263. [[CrossRef](#)] [[PubMed](#)]
- Nofech-Mozes, S.; Vella, E.T.; Dhesy-Thind, S.; Hanna, W.M. Cancer care Ontario guideline recommendations for hormone receptor testing in breast cancer. *Clin. Oncol.* **2012**, *24*, 684–696. [[CrossRef](#)]
- Wolff, A.C.; Hammond, M.E.; Hayes, D.F. Re: Predictability of adjuvant trastuzumab benefit in N9831 patients using the ASCO/CAP HER2-positivity criteria. *J. Natl. Cancer Inst.* **2012**, *104*, 957–958. [[CrossRef](#)]

18. Iglesias, J.E.; Liu, C.Y.; Thompson, P.M.; Tu, Z. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Trans. Med. Imaging* **2011**, *30*, 1617–1634. [[CrossRef](#)]
19. Fonov, V.; Evans, A.C.; Botteron, K.; Almli, C.R.; McKinstry, R.C.; Collins, D.L.; Brain Development Cooperative Group. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* **2011**, *54*, 313–327. [[CrossRef](#)]
20. Modat, M.; Cash, D.M.; Daga, P.; Winston, G.P.; Duncan, J.S.; Ourselin, S. Global image registration using a symmetric block-matching approach. *J. Med. Imaging* **2014**, *1*, 024003. [[CrossRef](#)]
21. Tustison, N.J.; Avants, B.B.; Cook, P.A.; Zheng, Y.; Egan, A.; Yushkevich, P.A.; Gee, J.C. N4ITK: Improved N3 bias correction. *IEEE Trans. Med. Imaging* **2010**, *29*, 1310–1320. [[CrossRef](#)] [[PubMed](#)]
22. Reinhold, J.C.; Dewey, B.E.; Carass, A.; Prince, J.L. Evaluating the Impact of Intensity Normalization on MR Image Synthesis. In Proceedings of the SPIE Medical Imaging, San Diego, CA, USA, 16–21 February 2019; Volume 10949. [[CrossRef](#)]
23. van Griethuysen, J.J.M.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.H.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)] [[PubMed](#)]
24. Abraham, A.; Pedregosa, F.; Eickenberg, M.; Gervais, P.; Mueller, A.; Kossaiji, J.; Gramfort, A.; Thirion, B.; Varoquaux, G. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinform.* **2014**, *8*, 14. [[CrossRef](#)] [[PubMed](#)]
25. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
26. Sarica, A.; Cerasa, A.; Quattrone, A. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer’s Disease: A Systematic Review. *Front. Aging Neurosci.* **2017**, *9*, 329. [[CrossRef](#)] [[PubMed](#)]
27. Karpievitch, Y.V.; Hill, E.G.; Leclerc, A.P.; Dabney, A.R.; Almeida, J.S. An introspective comparison of random forest-based classifiers for the analysis of cluster-correlated data by way of RF++. *PLoS ONE* **2009**, *4*, e7087. [[CrossRef](#)]
28. Louppe, G. Understanding variable importances in forests of randomized trees. In Proceedings of the Advances in Neural Information Processing Systems 26, Lake Tahoe, NV, USA, 5–8 December 2013.
29. Gresser, E.; Schachtner, B.; Stuber, A.T.; Solyanik, O.; Schreier, A.; Huber, T.; Froelich, M.F.; Magistro, G.; Kretschmer, A.; Stief, C.; et al. Performance variability of radiomics machine learning models for the detection of clinically significant prostate cancer in heterogeneous MRI datasets. *Quant. Imaging Med. Surg.* **2022**, *12*, 4990–5003. [[CrossRef](#)]
30. Rigatti, S.J. Random forest. *J. Insur. Med.* **2017**, *47*, 31–39. [[CrossRef](#)]
31. Limkin, E.J.; Sun, R.; Derclé, L.; Zacharaki, E.I.; Robert, C.; Reuze, S.; Schernberg, A.; Paragios, N.; Deutsch, E.; Ferte, C. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann. Oncol.* **2017**, *28*, 1191–1206. [[CrossRef](#)]
32. Reps, J.M.; Ryan, P.; Rijnbeek, P.R. Investigating the impact of development and internal validation design when training prognostic models using a retrospective cohort in big US observational healthcare data. *BMJ Open* **2021**, *11*, e050146. [[CrossRef](#)]
33. Mason, S.J.; Graham, N.E. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Q. J. R. Meteorol. Soc.* **2002**, *128*, 2145–2166. [[CrossRef](#)]
34. Matthews, B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta* **1975**, *405*, 442–451. [[CrossRef](#)] [[PubMed](#)]
35. Powers, D.M. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv* **2020**, arXiv:2010.16061.
36. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.C.; Muller, M. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77. [[CrossRef](#)]
37. Senkus, E.; Kyriakides, S.; Ohno, S.; Penault-Llorca, F.; Poortmans, P.; Rutgers, E.; Zackrisson, S.; Cardoso, F.; Committee, E.G. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2015**, *26* (Suppl. 5), v8–v30. [[CrossRef](#)]
38. Lev, S. Targeted therapy and drug resistance in triple-negative breast cancer: The EGFR axis. *Biochem. Soc. Trans.* **2020**, *48*, 657–665. [[CrossRef](#)]
39. Carlson, R.W.; Allred, D.C.; Anderson, B.O.; Burstein, H.J.; Edge, S.B.; Farrar, W.B.; Forero, A.; Giordano, S.H.; Goldstein, L.J.; Gradishar, W.J.; et al. Metastatic breast cancer, version 1.2012: Featured updates to the NCCN guidelines. *J. Natl. Compr. Cancer Netw.* **2012**, *10*, 821–829. [[CrossRef](#)]
40. Catalano, O.A.; Horn, G.L.; Signore, A.; Iannace, C.; Lepore, M.; Vangel, M.; Luongo, A.; Catalano, M.; Lehman, C.; Salvatore, M.; et al. PET/MR in invasive ductal breast cancer: Correlation between imaging markers and histological phenotype. *Br. J. Cancer* **2017**, *116*, 893–902. [[CrossRef](#)]
41. Incoronato, M.; Grimaldi, A.M.; Cavaliere, C.; Inglese, M.; Mirabelli, P.; Monti, S.; Ferbo, U.; Nicolai, E.; Soricelli, A.; Catalano, O.A.; et al. Relationship between functional imaging and immunohistochemical markers and prediction of breast cancer subtype: A PET/MRI study. *Eur. J. Nucl. Med. Mol. Imaging* **2018**, *45*, 1680–1693. [[CrossRef](#)]
42. Saha, A.; Harowicz, M.R.; Grimm, L.J.; Kim, C.E.; Ghate, S.V.; Walsh, R.; Mazurowski, M.A. A machine learning approach to radiogenomics of breast cancer: A study of 922 subjects and 529 DCE-MRI features. *Br. J. Cancer* **2018**, *119*, 508–516. [[CrossRef](#)]
43. Luo, X.; Xie, H.; Yang, Y.; Zhang, C.; Zhang, Y.; Li, Y.; Yang, Q.; Wang, D.; Luo, Y.; Mai, Z.; et al. Radiomic Signatures for Predicting Receptor Status in Breast Cancer Brain Metastases. *Front. Oncol.* **2022**, *12*, 878388. [[CrossRef](#)]
44. Cao, R.; Pang, Z.; Wang, X.; Du, Z.; Chen, H.; Liu, J.; Yue, Z.; Wang, H.; Luo, Y.; Jiang, X. Radiomics evaluates the EGFR mutation status from the brain metastasis: A multi-center study. *Phys. Med. Biol.* **2022**, *67*, 125003. [[CrossRef](#)]

45. Hohm, A.; Karremann, M.; Gielen, G.H.; Pietsch, T.; Warmuth-Metz, M.; Vandergrift, L.A.; Bison, B.; Stock, A.; Hoffmann, M.; Pham, M.; et al. Magnetic Resonance Imaging Characteristics of Molecular Subgroups in Pediatric H3 K27M Mutant Diffuse Midline Glioma. *Clin. Neuroradiol.* **2022**, *32*, 249–258. [\[CrossRef\]](#)
46. Krawczyk, N.; Fehm, T.; Banyas-Paluchowski, M.; Janni, W.; Schramm, A. Liquid Biopsy in Metastasized Breast Cancer as Basis for Treatment Decisions. *Oncol. Res. Treat.* **2016**, *39*, 112–116. [\[CrossRef\]](#)
47. Foda, A.; Kellner, E.; Gunawardana, A.; Gao, X.; Janz, M.; Kufner, A.; Khalil, A.A.; Geran, R.; Mekle, R.; Fiebach, J.B.; et al. Differentiation of Cerebral Neoplasms with Vessel Size Imaging (VSI). *Clin. Neuroradiol.* **2022**, *32*, 239–248. [\[CrossRef\]](#)
48. Zhou, M.; Scott, J.; Chaudhury, B.; Hall, L.; Goldgof, D.; Yeom, K.W.; Iv, M.; Ou, Y.; Kalpathy-Cramer, J.; Napel, S.; et al. Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches. *Am. J. Neuroradiol.* **2018**, *39*, 208–216. [\[CrossRef\]](#)
49. Cho, N. Molecular subtypes and imaging phenotypes of breast cancer. *Ultrasonography* **2016**, *35*, 281–288. [\[CrossRef\]](#)
50. Youk, J.H.; Son, E.J.; Chung, J.; Kim, J.A.; Kim, E.K. Triple-negative invasive breast cancer on dynamic contrast-enhanced and diffusion-weighted MR imaging: Comparison with other breast cancer subtypes. *Eur. Radiol.* **2012**, *22*, 1724–1734. [\[CrossRef\]](#)
51. Aerts, H.J. The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review. *JAMA Oncol.* **2016**, *2*, 1636–1642. [\[CrossRef\]](#)
52. Lambin, P.; Leijenaar, R.T.; Deist, T.M.; Peerlings, J.; De Jong, E.E.; Van Timmeren, J.; Sanduleanu, S.; Larue, R.T.; Even, A.J.; Jochems, A. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [\[CrossRef\]](#)
53. Gillies, R.; Kinahan, P.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [\[CrossRef\]](#)
54. Kniep, H.C.; Madesta, F.; Schneider, T.; Hanning, U.; Schonfeld, M.H.; Schon, G.; Fiehler, J.; Gauer, T.; Werner, R.; Gellissen, S. Radiomics of Brain MRI: Utility in Prediction of Metastatic Tumor Type. *Radiology* **2019**, *290*, 479–487. [\[CrossRef\]](#)
55. Coates, A.S.; Winer, E.P.; Goldhirsch, A.; Gelber, R.D.; Gnant, M.; Piccart-Gebhart, M.; Thürlimann, B.; Senn, H.-J.; Members, P.; André, F. Tailoring therapies—Improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann. Oncol.* **2015**, *26*, 1533–1546. [\[CrossRef\]](#)
56. Parmar, C.; Rios Velazquez, E.; Leijenaar, R.; Jermoumi, M.; Carvalho, S.; Mak, R.H.; Mitra, S.; Shankar, B.U.; Kikinis, R.; Haibe-Kains, B. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS ONE* **2014**, *9*, e102107. [\[CrossRef\]](#)
57. Yip, S.S.; Aerts, H.J. Applications and limitations of radiomics. *Phys. Med. Biol.* **2016**, *61*, R150. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Diskordante und konvertierende Rezeptorexpressionen in Hirnmetastasen von Brustkrebs: MRT Radiomics-basierte nicht-invasive Bestimmung des Rezeptorstatus

2 Einleitung

2.1 Radiomics und Machine Learning in der medizinischen Bilddatenanalyse

Radiomics ist ein sich rasch entwickelndes Forschungsgebiet, das sich mit der Extraktion quantitativer Metriken - den so genannten radiomischen Merkmalen - aus medizinischen Bildern befasst. Unter medizinischen Bildern wird hier vor allem die CT-, MRT- und PET-Bildgebung verstanden, welche standardmäßig im klinischen Alltag erhoben wird (Guiot, Vaidyanathan et al. 2022). Die Anwendungsgebiete für Radiomics sind vielfältig und umfassen verschiedene Bereiche der medizinischen Bildgebung. Beispielsweise kann Radiomics in der Krebserkennung- und Charakterisierung, in der Therapieplanung und Überwachung, zur Prognose- und Risikobewertung und als prädiktive Diagnostik oder zur Identifizierung bildbasierter Biomarker eingesetzt werden (Gillies, Kinahan et al. 2016). Während in der konventionellen radiologischen Befundung vorwiegend semantische Merkmale wie Form, Kontrastmittelverhalten und Lokalisation einer Läsion betrachtet werden und es inter- und intra-Rater spezifische Variabilität gibt, werden in radiomischen Merkmalen quantitative Gewebe- und Läsionseigenschaften wie Heterogenität und Form erfasst. Diese können allein oder in Kombination mit demografischen, histologischen, genomischen oder proteomischen Daten zur Lösung klinischer Fragestellungen verwendet werden (Rios Velazquez, Parmar et al. 2017, Kniep, Madesta et al. 2019, Mayerhoefer, Materka et al. 2020). Obwohl der Begriff Radiomics nicht streng definiert ist, zielt Radiomics im Allgemeinen darauf ab, quantitative und idealerweise reproduzierbare Informationen aus diagnostischen Bildern zu extrahieren, einschließlich komplexer Muster, die für das menschliche Auge schwer zu erkennen oder zu quantifizieren sind (Aerts 2016, Gillies, Kinahan et al. 2016). Es wird daher versucht, die gegebenen Bildinformationen möglichst vollständig, quantitativ und standardisiert zu erfassen. Die extrahierten radiomischen Bildeigenschaften lassen sich grundsätzlich in 3 Hauptmerkmalsklassen einteilen: Form, statistische Verteilung der Grauwerte und Textur (Haralick, Shanmugam et al. 1973). Die erhobenen Bildinformationen werden anschließend maschinellen Lernalgorithmen (ML) in Form von Trainings-/Validierungs- und Testdatensätzen zugeführt. Ein Modell, das diskrete Kategorien (manchmal auch als Klassen bezeichnet) erzeugt, wird als Klassifizierungsalgorithmus bezeichnet. Beispiele für Klassifizierungsalgorithmen sind solche, die vorhersagen, ob ein Tumor gutartig oder bösartig ist. Die zugrunde liegende Tumorbiologie kann die physikalischen Eigenschaften des Gewebes verändern, was sich auf die Muster der Bildpixel und somit auf die radiometrischen Merkmale auswirkt (Abbasian Ardakani, Bureau et al. 2022). Ein Modell, das eine Vorhersage für einen kontinuierlichen Wert liefert, wird als Regressionsalgorithmus bezeichnet. Im ML könnte ein Algorithmus, der als

Regressionsalgorithmus bezeichnet wird, verwendet werden, um die Lebenserwartung einer Person oder die verträgliche Dosis einer Chemotherapie vorherzusagen (Sidey-Gibbons and Sidey-Gibbons 2019). Im Rahmen dieser Promotionsarbeit wurde das Potential einer Radiomics-basierten Klassifikation des Rezeptorstatus von zerebralen Metastasen des Mammakarzinoms mittels Random Forest (RF) Modellen analysiert. RF Modelle können sowohl als Klassifikations- als auch Regressionsverfahren konzipiert sein. Im Fall dieser Arbeit lieferten die RFs zwar zunächst kontinuierliche Werte in Form von Wahrscheinlichkeiten als Output, diese wurden jedoch anschließend mithilfe von Thresholds in Klassen binarisiert. RF Klassifikatoren haben den Vorteil, dass sie nicht zum „Overfitting“ neigen und dass sie viele heterogene Prädiktoren miteinschließen können (Breiman 2001, Sarica, Cerasa et al. 2017). Des Weiteren sind sie robust gegen cluster-korrelierte Daten (Karpievitch, Hill et al. 2009), weswegen sich zum Beispiel wie in der vorliegenden Dissertation gut mehrere Metastasen eines einzelnen Patienten in die Analyse miteinbinden lassen. RF Modelle können sowohl im „Supervised Learning“ als auch im „Unsupervised Learning“ eingesetzt werden. Radiomics und ML können also für die Extraktion und Kategorisierung von großen Datenmengen für prädiktive oder prognostische Fragestellungen benutzt werden. In einer Studie konnte gezeigt werden, dass mithilfe einiger extrahierter Bildfeatures über 80% des Genexpressionsmusters eines hepatozellulären Karzinoms vorhergesagt werden konnten (Kumar, Gu et al. 2012). Dieser sich entwickelnde Zweig von Radiomics, der bildgebende Merkmale mit der Genexpression verknüpft, ist heute als Radiogenomik bekannt (Lo Gullo, Daimiel et al. 2020).

Trotz der rasant zunehmenden Forschung und dem großen Potential von Radiomics und ML bestehen einige Herausforderungen. Als häufiger Kritikpunkt wird die mögliche Instabilität einiger extrahierter Bildmerkmale angesehen, welche von vielen Faktoren abhängig ist (Scanner-Modell, Scan-Protokoll, Patientenlagerung, verwendetes Kontrastmittel) (Galavis, Hollensen et al. 2010, Yang, Wu et al. 2022). Hinzu kommt, dass sich diese Faktoren oft von Klinik zu Klinik unterscheiden. Ein weiterer regelmäßig angeführter Kritikpunkt ist, dass Modelle, die auf sehr heterogenen Daten trainiert wurden, aufgrund von möglichen patientenspezifischen Confoundern (z.B. Alter, Geschlecht) schlechtere Prädiktionsgüten aufweisen als Modelle, die auf einem homogenen, stratifizierten Patientenkollektiv trainiert wurden (Hu, Yoon et al. 2019). Eine entsprechende Strategie könnte also sein, individuelle Modelle für bestimmte Patientenkollektive zu etablieren, um eine bestmögliche Prädiktion für den einzelnen Patienten zu erzielen (Kniep 2020). Zudem bewerten viele Untersuchungen hauptsächlich die Leistung von Prädiktionsmodellen auf Kohortenebene. Für den Einsatz in der klinischen Praxis ist es jedoch auch wichtig, die Unsicherheit auf der Ebene des einzelnen Patienten zu bewerten und zu quantifizieren (Hu and Swanson 2020, Kniep 2020).

Das Ziel von Radiomics und ML Analysen, die sich auf Metastasen konzentrieren, besteht regelmäßig darin, das Therapieansprechen zu erkennen, den Unterschied zwischen Pseudoprogress und echtem Progress festzustellen, den Primarius einer Metastase zu identifizieren oder Gewebe hinsichtlich Mutationen oder Rezeptorstatus zu charakterisieren (Bhatia, Birger et al. 2019, Kniep, Madesta et al. 2019, Basler, Gabryś et al. 2020, Luo, Xie et al. 2022).

Trotz der aktuell noch bestehenden Herausforderungen hinsichtlich Reproduzierbarkeit, Standardisierung und Nachvollziehbarkeit weisen Radiomics und ML erhebliches Potential auf, weswegen weitere Forschung geleistet werden sollte, um diese Limitationen zu bewältigen und die Techniken im klinischen Alltag zu etablieren (Scapicchio, Gabelloni et al. 2021).

2.2 Hirnmetastasen von Brustkrebs: Klinische Implikationen

Das Auftreten von Hirnmetastasen hat erhebliche Auswirkungen auf die Prognose und wirkt sich auf die gesundheitsbezogene Lebensqualität bei Patienten und Patientinnen mit Brustkrebs aus (Cardoso, Paluch-Shimon et al. 2020, Kanchan, Siddiqui et al. 2020, Watase, Shiino et al. 2021). Bei Brustkrebs liegt die gemeldete Inzidenz für zerebrale Metastasen zwischen 10-30 % (Chamberlain, Baik et al. 2017, Hosonaga, Saya et al. 2020, Wang, Ye et al. 2021). Bei der Brustkrebstherapie wird der Status des Östrogenrezeptors-alpha (ER), des Progesteronrezeptors (PR) und des humanen epidermalen Wachstumsfaktor-Rezeptor 2 (HER2) des Primärtumors zu den wichtigsten Determinanten für gezielte Behandlungsschemata gezählt (Gennari, Andre et al. 2021). Als gezielte Therapien werden unter anderem Aromatasehemmer, Östrogenrezeptormodulatoren, CDK4/6-Inhibitoren und monoklonale Antikörper gegen HER2 oder gegen den vaskulären endothelialen Wachstumsfaktor (VEGF) eingesetzt. Im Falle einer metastasierten Erkrankung basieren therapeutische Entscheidungen häufig auf den Rezeptorexpressionen des Primärtumors. Resektionen und Biopsien von Hirnmetastasen können nur dann in Betracht gezogen werden, wenn die Lage der Metastasen, die Anzahl und der allgemeine Gesundheitszustand des Patienten diese invasiven Verfahren zulassen (Jung, Lee et al. 2018, Ute-Susann Albert 2021). Diesbezüglich haben mehrere Studien gezeigt, dass Metastasen eine beträchtliche Tendenz zu Mutationen in ihrer Rezeptorexpression aufweisen, was häufig zu einem abweichendem Rezeptorstatus im Vergleich zum Primarius führt (Aurilio, Monfardini et al. 2013, Hoefnagel, van der Groep et al. 2013, Kroigard, Larsen et al. 2016, Schrijver, Selenica et al. 2018, Schrijver, Suijkerbuijk et al. 2018, Yang, Forbes et al. 2018, Michel, Oppong et al. 2021, Yi, Yu et al. 2021, Zhao, Sun et al. 2021). Eine durchgeführte Meta-Analyse ergab folgende Raten für die Rezeptorkonversion bei Patientinnen mit metastasiertem Brustkrebs gegenüber der Rezeptorexpression des Primärtumors: Für ER, PR und HER2 lagen die gepoolten Konversionsraten von positiv zu negativ bei 22,5 %, 49,4 % bzw. 21,3 %; die Konversionsraten von negativ zu positiv betragen 21,5 %, 15,9 % und 9,5 % (Schrijver, Suijkerbuijk et al. 2018). Im Anschluss an diese Ergebnisse wurde in mehreren Leitlinien empfohlen, Metastasen zu biopsieren, um die Rezeptorexpression der Metastasen zu bestätigen (Gennari, Andre et al. 2021). Jüngste Aktualisierungen erweitern diesen Ansatz und empfehlen auch eine erneute Bewertung der metastatischen Rezeptorexpressionen im Verlauf, wenn dies klinisch möglich ist (Cardoso, Paluch-Shimon et al. 2020). Bei zerebralen Metastasen sind Biopsien jedoch mit dem Risiko von Blutungen, Infektionen und neurologischen Komplikationen verbunden oder sind technisch nicht durchführbar (Schrijver, Suijkerbuijk et al. 2018). Diesen Patienten könnte in Zukunft die nicht-invasive Bestimmung des Rezeptorstatus mithilfe fortschrittlicher radiologischer Bildgebung und

Auswertungstechniken helfen. Ein solcher Ansatz könnte eine kontinuierliche Überwachung der Rezeptorexpression und eine dynamische Anpassung der angewandten gezielten Behandlungsoptionen ermöglichen. Eine Bestimmung und fortlaufende Überwachung des Rezeptorstatus der Metastasen wäre so zu niedrigeren Kosten und mit einem geringen Risiko möglich. Die Metastase kann ganzheitlicher als bei der Biopsie bewertet werden und schnellere Kontrollen wären möglich (Abbasian Ardakani, Bureau et al. 2022).

In der vorliegenden explorativen Studie wird das Potenzial der Bestimmung des Rezeptorstatus mit Hilfe eines maschinellen Lernansatzes untersucht. Wir stellen die Hypothese auf, dass quantitative bildgebende Biomarker von Hirnmetastasen, die aus multispektralen Standard-MR-Bilddaten extrahiert wurden, zur nicht-invasiven Bestimmung des Rezeptorstatus dieser metastatischen Läsionen verwendet werden können.

3 Material und Methoden

3.1 Patientenkollektiv

Die retrospektiv durchgeführte Einzel-Center-Studie wurde von der lokalen Ethikkommission genehmigt und es war aufgrund der Anonymisierung der Studiendaten keine Patientenzustimmung erforderlich (WF-144/20). Die Studie wurde in Übereinstimmung mit der Deklaration von Helsinki durchgeführt. Eine Übersicht der in dieser Arbeit durchgeführten Teilschritte zur Bestimmung des Rezeptorstatus der Hirnmetastasen ist in Abbildung 1 gegeben.

In die Studie eingeschlossen wurden 412 Hirnmetastasen von 106 Brustkrebspatienten und Brustkrebspatientinnen, die zwischen September 2007 und September 2021 in der Strahlentherapie des Universitätsklinikums Hamburg-Eppendorf vorstellig wurden. Bei den eingeschlossenen Patienten und Patientinnen wurde der Rezeptorstatus des Tumors durch Biopsie des Primarius und anschließende histologische Aufarbeitung gesichert, in einigen Fällen wurden auch die Metastasen selbst biopsiert und der Rezeptorstatus bestimmt. Von der Analyse ausgeschlossen wurden Metastasen, die bereits vorbehandelt wurden (Bestrahlung, Radiochirurgie, Operation), da die Vorbehandlung potenziell zu einer Veränderung der radiomischen Signatur führt. Eine Übersicht des Patientenkollektivs ist in Tabelle 1 dargestellt.

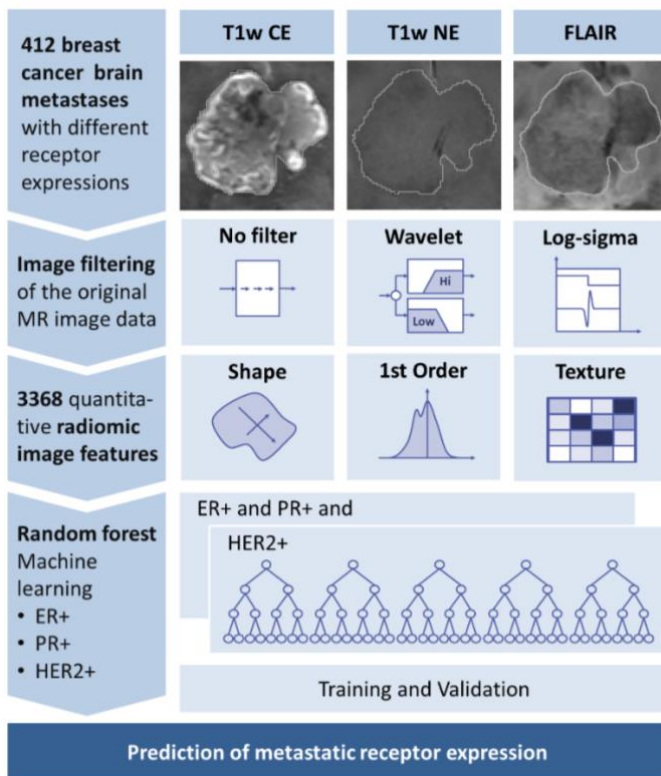


Abbildung 1: Übersicht des durchgeführten Algorithmus für die Rezeptorstatus-Vorhersage. T1w CE, kontrastverstärkte T1-gewichtete MR-Bilder; T1w NE, nicht kontrastverstärkte T1-gewichtete MR-Bilder, FLAIR, flüssigkeitsabgeschwächte Inversionserholung; ER, Östrogenrezeptor; PR, Progesteronrezeptor; HER2, humaner epidermaler Wachstumsfaktorrezeptor 2.

	Anzahl an		Alter der Stichprobenpatienten			Anzahl Metastasen der Stichprobenpatienten		
	Metastasen	Patienten	Mean	Min	Max	Mean	Min	Max
Alle Patienten	412	106 30* 76**	55.9	33	84	5.4	1	66
Classifier 1 (ER)	<i>Total Samples</i>							
ER+	190	53	58.3	33	76	4	1	63
ER-	222	53	53.9	36	79	5.9	1	66
Classifier 2 (PR)	<i>Total Samples</i>							
PR+	147	37	59.5	39	76	4.5	1	63
PR-	265	69	53.9	33	84	5.7	1	66
Classifier 3 (HER)	<i>Total Samples</i>							
HER+	169	45	54.5	36	81	4.1	1	63
HER-	243	61	56.9	33	84	6.1	1	66

Tabelle 1: Demographische and klinische Charakteristiken der eingeschlossenen Metastasen. Die Tabelle zeigt die Anzahl eingeschlossener Metastasen, die Anzahl der Patienten, das durchschnittliche Alter bei Diagnose der Hirnmetastasen geordnet nach Rezeptorstatus und die Anzahl der Hirnmetastasen des durchschnittlichen Patienten geordnet nach Rezeptorstatus. *Anzahl eingeschlossener Patienten mit einer einzelnen Hirnmetastase. ** Anzahl eingeschlossener Patienten mit multiplen Hirnmetastasen.

3.2 Bildgebung, Segmentierung, Bildpostprozessierung

Die Bildgebung erfolgte durch 1,5 Tesla-MRT-Scanner (Magnetom® Symphony, Magnetom® Avanto, and Magnetom® Sonata, hergestellt von Siemens Healthcare, Erlangen, Deutschland) oder durch 3 Tesla-MRT-Scanner (Ingenia®, Philips Medical Systems, Best, Niederlande; Magnetom® Skyra and Magnetom® Vida, Siemens Healthineers AG, Erlangen, Deutschland). Die Bildgebungsverfahren umfassten eine axiale Fluid Attenuated Inversion Recovery (FLAIR) Sequenz ohne Kontrastmittel und eine axiale T1-gewichtete Spin-Echo-Sequenz mit Flusskompensation (T1w nativ), ebenfalls ohne Kontrastmittel. Nach Verabreichung eines gewichtsadjustierten Gadolinium-basierten Kontrastmittels (0,1 mmol/kg Körpergewicht) wurden entweder kontrastmittelverstärkte axiale T1-gewichtete Spin-Echo-Sequenzen mit Flusskompensation oder dreidimensionale T1-gewichtete Gradienten-Echo-Sequenzen aufgenommen (T1w KM). Die genauen Sequenzparameter variierten je nach MRT-Gerät und spiegeln die Vielfalt der Bilddaten in der klinischen Praxis wider.

Die Segmentierungen der Hirnmetastasen wurden anhand der kontrastverstärkten T1-Bilder durchgeführt. Hierfür wurde ein semi-automatisierter Ansatz auf Basis der ITK-SNAP-Software (Biomedical Imaging Resource; Penn Image Computing and Science Laboratory University of Pennsylvania (PA), Scientific Computing and Imaging Institute of Utah (UT)) verwendet. Kontrolliert wurden die Segmentierungen von zwei zu allen klinischen Daten verblindeten Neuroradiologen. Die regions of interest (ROIs) wurden nach dem sichtbaren gross tumor volume (GTV) definiert. Bei Inkonsistenzen hinsichtlich der Segmentierungen wurden die sich überlappenden ROIs gebildet und als finale Konsensus-ROIs verwendet.

Im Rahmen der Postprozessierung wurde durch die ROBEX Software (NITRC; University of Massachusetts Medical School in Worcester, MA) (Iglesias, Liu et al. 2011) die Knochenmasse auf den Bildern entfernt und alle Bilder sowie ROIs wurden mit NiftyReg (Centre for Medical Image Computing; University College London, UK) auf gemeinsame standardisierte MNI152 Atlanten registriert (Fonov, Evans et al. 2011). Anschließend wurden die Bilddaten N4ITK-verzerrungskorrigiert (Tustison, Avants et al. 2010) und mittels „piecewise linear histogram matching“ intensitätsnormalisiert (Reinhold, Dewey et al. 2019). Diese Schritte dienten der Standardisierung der Bilder und somit auch der radiomischen Bildmerkmale. Die erfolgreiche Bildpostprozessierung wurde visuell durch zwei Neuroradiologen bestätigt.

3.3 Extraktion der quantitativen Bildmerkmale und Machine Learning

Die Merkmalsextraktion der MRT-Bilder erfolgte durch das PyRadiomics Python Paket Version 2.2.0 (van Griethuysen, Fedorov et al. 2017). Um sinnvolle Ergebnisse hinsichtlich der Merkmalsausprägungen zu erhalten, wurden Kriterien für die Mindestgröße einer Metastase festgelegt, um sie in die Analyse einzuschließen. Bei Patienten und Patientinnen, die mehr als 10 Hirnmetastasen aufwiesen, wurden nur die 10 größten Hirnmetastasen eingeschlossen, um cluster-korrelierte Daten zu vermeiden. Insgesamt wurden 3368 Bildmerkmale (erster Ordnung/statistische Verteilung der Grauwerte, Textur, Form) für jede Metastase extrahiert und gemeinsam mit dem Patientenalter als Input-Daten in dem Machine Learning Algorithmus verwendet.

Die ML-basierte Klassifikation wurde mit der Python Scikit-Learn Umgebung implementiert (Abraham, Pedregosa et al. 2014). Hierbei handelt es sich um eine Open-Source Python Bibliothek für maschinelles Lernen und zur prädiktiven Datenanalyse. Zur Klassifikation wurde ein Ensemble-begründeter Random Forest (RF) benutzt. RF Algorithmen weisen einige Vorteile gegenüber anderen Ansätzen auf: (a) RF Algorithmen verfügen über einfacher nachvollziehbare Entscheidungswege und sind keine schwer zu entschlüsselnden „Black-Boxen“ wie neuronale Netze (Zhang, Beck et al. 2018). (b) Der Algorithmus erfordert keine Normalisierung der unabhängigen Variablen (Kniep 2020). (c) Aufgrund des Gesetzes der großen Zahlen neigen sie nicht zum „Overfitting“ und können viele heterogene Prädiktoren miteinschließen (Breiman 2001, Sarica, Cerasa et al. 2017). (d) Sie sind relativ robust gegenüber Ausreißern und Rauschen. (e) Sie liefern nützliche interne Schätzungen zu Fehler, Stärke, Korrelation und Bedeutung der Variablen (Breiman 2001). Da sich in unseren Datensätzen multiple Metastasen einzelner Patienten befinden, kann dies potenziell zu cluster-korrelierten Überschätzungen der Ergebnisse führen. Von Karpievitch et al. wurde gezeigt, dass RF Algorithmen weniger anfällig gegenüber cluster-korrelierten Daten sind (Karpievitch, Hill et al. 2009).

In unserer Studie wurden drei Klassifikatoren trainiert: Klassifikator 1, für die Unterscheidung von Östrogenrezeptor-positiven versus -negativen Metastasen; Klassifikator 2 für die Unterscheidung von Progesteronrezeptor-positiven versus -negativen Metastasen; Klassifikator 3 für die Unterscheidung von humanen epidermalen Wachstumsfaktor-2-Rezeptor-positiven versus -negativen Metastasen.

3.4 Selektion quantitativer Bildmerkmale, Modellhyperparameter

Die Vorhersagekraft der verwendeten Bildmerkmale wurde für jeden Trainingsdatensatz separat auf der Grundlage von Gini-Verunreinigungsmaßen berechnet (Louppe 2013). Die Gini-Verunreinigung zeigt, wie gut ein Merkmal die Gesamtstichproben der binären Klassen in einem bestimmten Knotenpunkt aufteilt (Disha and Waheed 2022). Da zu erwarten ist, dass viele Merkmale korreliert sind, zum Beispiel aufgrund der Verwendung mehrerer Filter auf demselben Bild, und aufgrund des Risikos einer Überanpassung des Modells mit zu vielen Merkmalen, beschränkten wir die Anzahl der verwendeten Merkmale auf die 30 wichtigsten Merkmale des jeweiligen Klassifikators (Gennari, Andre et al. 2021).

Bei allen Evaluierungsläufen der Modelle wurde die Bedeutung der Merkmale nur auf der Grundlage der jeweiligen Trainingsdaten bestimmt.

Die ensemblebasierten Random-Forest-Algorithmen, erstellen eine festgelegte Anzahl von Entscheidungsbäumen. Jeder Baum wird auf der Grundlage von Teilproben mit Ersatz (Bootstrapping) trainiert (Breiman 2001). Um sicherzustellen, dass der vorgeschlagene Algorithmus weitreichend anwendbar ist und „Overfitting“ sowohl in Bezug auf die Parameterabstimmung als auch im Allgemeinen vermieden wird, wurden die Einstellungen der Random Forest Klassifikatoren von Anfang an gemäß gängigen Standardparameterwerten festgelegt. So wurde die Anzahl der Bäume entsprechend des Scikit-Learn Random Forest Defaults auf 1000 begrenzt. Die Anzahl der Entscheidungsmerkmale je Baum entspricht der Quadratwurzel aller extrahierten Bildmerkmale. Diese Konvention hat sich in der Literatur als gängige Praxis etabliert. Die finale Klassifizierungsleistung wird jedoch insgesamt nur geringfügig von der Hyperparameteroptimierung beeinflusst beziehungsweise befinden sich mit den Standardeinstellungen in den meisten Fällen nahe des Optimums (Bernard, Heutte et al. 2009, Probst, Bischl et al. 2018).

3.5 Erstellung der Trainings- und Testsets

Für das Training der Klassifikatoren sowie für die statistische Auswertung wurden aus den 412 Metastasen zehn stratifizierte 5-fache Kreuzvalidierungssets gebildet, die zufällig ausgewählt wurden. Die Kreuzvalidierung ist ein Ansatz zur mehrfachen Nutzung von Trainings- und Genauigkeitsbewertungsstichproben, wodurch die Zuverlässigkeit der Ergebnisse verbessert werden kann (A. Ramezan, A. Warner et al. 2019). Um eine mögliche Überschätzung der Klassifikationsleistung bei cluster-korrelierten Daten zu vermeiden, wurden für Patienten mit mehreren Metastasen die quantitativen Bildmerkmale entweder vollständig dem Test- oder dem Trainingsdatensatz zugeordnet. Die Performance des Modells wurde ausschließlich anhand der unabhängigen Validierungs-/Testdatensätze evaluiert.

3.6 Statistik

Um die Performance-Leistung der verschiedenen Klassifikatoren (ER, PR, HER2) bezogen auf die Validierungs-/Testsets zu beurteilen, wurde auf die Receiver-Operating-Characteristic Areas-under-the-Curve (ROC AUCs) zurückgegriffen. Eine ROC-Kurve ist eine grafische Darstellung der Leistungsfähigkeit eines binären Klassifizierungsmodells für alle Klassifizierungsschwellenwerte. Die P-Werte wurden unter Verwendung des Verification R-Pakets Version 1.42 basierend auf der Mann-Whitney/Wilcoxon U-Statistik berechnet. Statistische Signifikanz wurde bei P-Werten $<0,05$ angenommen. In den Berechnungen wurde die Nullhypothese $H_0: AUC = 0,5$ berücksichtigt, was bedeutet, dass die Vorhersage der Klassifikatoren zufällig wäre (Mason and Graham 2002).

Klassifizierungsinstabilitäten wurden durch die 10-fache Ziehung von 5 Kreuzvalidierungssätzen evaluiert.

Weiterhin wurde der maximale Matthew Korrelationskoeffizient (MCC) berechnet und die ROC-Kurven an diesen Punkten analysiert. Der MCC berücksichtigt alle Felder der Konfusionsmatrix und darüber hinaus den positiven und negativen prädiktiven Wert (Chicco and Jurman 2023). Er wird daher für unvoreingenommene Vergleiche von binären Klassifikatoren favorisiert (Powers 2020).

Konfidenzintervalle (CIs) für Sensitivitäten und Spezifitäten wurden mithilfe des pROC-R-Pakets Version 1.18.0 sowie der Python Scikit-Learn Metrics Umgebung Version 1.1.1 bestimmt (Robin, Turck et al. 2011, Abraham, Pedregosa et al. 2014).

4 Ergebnisse

4.1 Studienkollektiv

Das Studienkollektiv umfasste T1 kontrastverstärkte, T1 native und FLAIR MRT-Sequenzen von 412 Metastasen von 106 Patienten und Patientinnen mit Hirnmetastasen durch primären Brustkrebs. Das Trainingsset schloss 329 Samples, das unabhängige Validierungs-/Testset schloss 83 Samples in einem von zehn zufälligen Kreuzvalidierungs-Testdurchläufen ein.

4.2 Östrogenrezeptor

Die ROC-AUC in den Validierungssets für die Vorhersage des Rezeptorstatus auf der Grundlage multiparametrischer MR-Scans betrug 0,82 mit 95 % CI [0,78; 0,85]. Beim maximalen MCC-Cut-off-Punkt ergab der Klassifikator eine Spezifität von 68 % (95 % CI [65 %; 71 %]) bei einer Sensitivität von 84 % (95 % CI [81 %; 88 %]). Darüber hinaus lagen der positive prädiktive Wert (PPV) und der negative prädiktive Wert (NPV) bei 69 % bzw. 84 %, die Genauigkeit bei 76 % und der MCC bei 0,53 (Tabelle 2).

4.3 Progesteronrezeptor

Die ROC-AUC der Validierungssets lag bei 0,73 mit 95% CI [0,69; 0,77]. Beim maximalen MCC-Cut-off-Punkt (MCC = 0,42) wurden Spezifitäten von 82 % (95 % CI [77 %; 85 %]) bei Sensitivitäten von 59 % (95 % CI [56 %; 64 %]) beobachtet. PPV und NPV lagen bei 69 % bzw. 74 %; die Genauigkeit betrug 72 % (Tabelle 2).

4.4 Humaner EGF-Rezeptor

Die ROC-AUC der Validierungssets lag bei 0,74 mit 95% CI [0,70; 0,78]. Beim maximalen MCC-Cut-off-Punkt (MCC = 0,47) wurden Spezifitäten von 84 % (95 % CI [81 %; 88 %]) bei Sensitivitäten von

61 % (95 % CI [57 %; 65 %]) beobachtet. Der PPV lag bei 73 %, der NPV bei 76 % und die Genauigkeit bei 75 % (Tabelle 2).

Die vergleichende Analyse von zufällig gezogenen 5-fachen Kreuzvalidierungssätzen in 10 Durchläufen bestätigte die geringe Variabilität der Ergebnisse, was auf eine stabile Modellleistung hindeutet. Die ROC-AUCs der unabhängigen Validierungs-/Testdatensätze lagen für alle drei Klassifikatoren bei > 0,5 mit P-Werten < 0,05. Die Nullhypothese einer zufälligen Vorhersage durch die Klassifikatoren konnte daher verworfen werden.

Matthews Korrelationskoeffizient (MCC) am maximalen Arbeitspunkt

Rezeptor Status	AUC [95% CI]	Sensitivität [95% CI]	Spezifität [95% CI]	PPV [95% CI]	NPV [95% CI]	Genauigkeit [95% CI]	MCC [95% CI]
<i>5-fache Kreuzvalidierung (n = 412 Samples)</i>							
ER+	0,82 [0,78; 0,85]	84% [81%; 88%]	68% [65%; 71%]	69% [67%; 72%]	84% [82%; 88%]	76% [73%; 78%]	0,53 [0,49; 0,58]
PR+	0,73 [0,69; 0,77]	59% [56%; 64%]	82% [77%; 85%]	69% [67%; 76%]	74% [70%; 76%]	72% [69%; 75%]	0,42 [0,37; 0,48]
HER2+	0,74 [0,70; 0,78]	61% [57%; 65%]	84% [81%; 88%]	73% [72%; 80%]	76% [74%; 78%]	75% [73%; 76%]	0,47 [0,45; 0,52]

Tabelle 2: Metriken zur Klassifizierungsleistung der unabhängigen Validierungssätze. Metriken der Klassifizierungsleistung der Klassifikatoren des maschinellen Lernens. Die Klassifikator-Metriken sind an den Cut-off-Punkten entsprechend dem maximalen Matthews-Korrelationskoeffizienten dargestellt. AUC, Fläche unter der Kurve; ER, Östrogenrezeptor; PR, Progesteronrezeptor; HER2, humaner epidermaler Wachstumsfaktor-Rezeptor 2; PPV, positiver prädiktiver Wert; NPV, negativer prädiktiver Wert; MCC, Matthews-Korrelationskoeffizient; CI, Konfidenzintervall.

4.5 Merkmalswichtigkeit und radiomische Signatur

Die Wichtigkeit einzelner radiomischer Merkmale wurden mit den Gini-Verunreinigungsmaßen bestimmt. Bei der Analyse wurden die 100 wichtigsten Merkmale für ER, PR und HER2 betrachtet. Unter den 100 wichtigsten Bildmerkmalen gab es für alle Klassifikatoren nur 2 Übereinstimmungen: „FLAIR Wavelet-HHL Firstorder Range“ und „T1CE Wavelet-HLL Firstorder Skewness“. Vergleicht man jedoch nur die Klassifikatoren für PR und HER2, stimmten 98 der 100 wichtigsten Bildmerkmale überein. Die Wichtigkeit der verschiedenen MRT-Sequenzen waren insgesamt ähnlich verteilt. In unseren Studiendaten schien für PR und HER2 die T1 nicht kontrastmittelverstärkten Sequenzen einen leicht stärkeren Beitrag zu leisten (36% der gesamten Vorhersageleistung) und für ER lieferten die FLAIR-Sequenzen den größten Beitrag zur Klassifikations-Performance (45%). Betrachtet man die

wichtigsten Merkmale nach eingesetztem Filter, so stellt sich der Wavelet-Filter in allen Klassifikatoren (ER, PR, HER2) als am Bedeutsamsten heraus (67% versus 57% versus 57% der Vorhersageleistung). Die Wavelet-Transformation teilt ein Signal oder ein Bild in verschiedene Skalen und Frequenzen auf. Im Gegensatz zur Fourier-Transformation, die das gesamte Signal im Frequenzbereich darstellt, analysiert die Wavelet-Transformation Signale sowohl im Zeitbereich als auch im Frequenzbereich. Bei den Signalen handelt es sich in unserem Fall um 2-D-Graustufenbilder, in denen der Zeitbereich die räumliche Lage eines Pixels und der Frequenzbereich die Intensitäts- oder Farbvariation um einen Pixel herum darstellt (Wang 2001). Dadurch ist sie besonders gut geeignet, um lokalisierte Informationen in einem Bild zu erfassen. Sieht man sich weiterhin die Merkmalsklassen an, kommt man in dieser Analyse zu dem Schluss, dass die Texturmerkmale den größten Einfluss gegenüber Merkmalen erster Ordnung oder Formmerkmalen (78% versus 75% versus 75%) in allen drei Klassifikatoren aufweisen. Beurteilt man die radiomische Signatur von ER-positiven und ER-negativen Metastasen wird deutlich, dass ER-positive Metastasen eine geringere Graustufenbetonung aufweisen. Außerdem deutet die Betrachtung der 10 wichtigsten Merkmale bei der ER-Klassifikation an, dass ER-positive Metastasen eine höhere Zonen-Entropie aufweisen. Die Zonen-Entropie misst die Unsicherheit/Zufälligkeit in der Verteilung der Zonengrößen und Graustufen. Ein höherer Wert weist auf eine größere Heterogenität in den Texturmustern hin. Die radiomischen Signaturen der 10 wichtigsten Merkmale für PR/HER2 bestätigen ähnliche Prädiktor-Verteilungen für PR und HER2 -Metastasen, aber das Gesamtbild der Verteilungen der einzelnen Prädiktoren ist im Vergleich zu den ER-Metastasen heterogener. Dennoch scheinen unterschiedliche Merkmale der Graustufenbetonung auch bei der PR-Vorhersage von Bedeutung zu sein. Die „Mean Absolute Deviation“, also der mittlere Abstand aller Intensitätswerte vom Mittelwert des Bildfeldes, weist für PR-positive und HER2-positive Metastasen wesentlich höhere Werte auf als für PR-negative und HER2-negative Metastasen. Dennoch zeichnet sich dieser Prädiktor in beiden Fällen durch eine recht hohe Variabilität aus. Alle Unterschiede der mittleren Merkmalswerte zwischen rezeptorpositiven und -negativen Metastasen waren statistisch signifikant (P-Wert < 0,05).

5 Diskussion

Ziel der Arbeit war die Untersuchung des Potentials, quantitative Bildmerkmale aus standardmäßigen multiparametrischen MR-Scans zu verwenden, um den Rezeptorstatus von Hirnmetastasen bei Patienten und Patientinnen mit metastasiertem Brustkrebs nicht-invasiv vorherzusagen. Unsere Ergebnisse zeigen, dass Algorithmen des maschinellen Lernens den Rezeptorstatus prädiktieren können und AUCs von 0,82 für ER+ -Metastasen, 0,73 für PR+ -Metastasen und 0,74 für HER2+ -Metastasen erreichen. Die Studie zeigt, dass ein Ansatz mit künstlicher Intelligenz das Potenzial hat, die Rezeptorexpressionen bei Patienten mit metastasiertem Brustkrebs nicht-invasiv zu bestimmen. Angesichts des verwendeten Kreuzvalidierungsansatzes und der Heterogenität der Bilddaten in Bezug auf Scanner-Modell,

Aufnahmeauflösung und Protokoll können wir davon ausgehen, dass die Ergebnisse gut verallgemeinerbar sind (Kniep 2020).

In den letzten Jahren haben die Fortschritte bei den zielgerichteten Behandlungsmöglichkeiten für Patienten und Patientinnen mit primärem und fortgeschrittenem Brustkrebs dazu geführt, dass die Expression des Rezeptorstatus des Primärtumors zu einer wichtigen Determinante der Therapieschemata geworden ist, was zu einer deutlichen Verbesserung der Behandlungsergebnisse geführt hat (Senkus, Kyriakides et al. 2015, Cardoso, Paluch-Shimon et al. 2020, Lev 2020). Mit folgenden beschriebenen Häufigkeiten werden jedoch Rezeptorstatuskonversionen beobachtet: 19,3 % für ER, 30,9 % für PR und 10,3 % für HER2 -Expressionen (Schrijver, Suijkerbuijk et al. 2018). Auch in unserer Subgruppenanalyse von Patienten und Patientinnen, für die die Rezeptorstatus von Primarius und Metastasen vorlagen, trat eine Konversion von mindestens einem Rezeptorstatus (ER, PR, HER2) bei 8 von 38 Patientinnen auf, was einer Konversionsrate von 21 % entspricht. Die Rezeptor-Divergenz könnte in zwei möglichen Szenarien erhebliche Auswirkungen haben: Im Falle von Mutationen, die zu einer Zunahme der Rezeptorexpression führen, werden den Patienten keine entsprechenden Medikamente verschrieben und sie werden nicht von wirksamen zielgerichteten Behandlungsoptionen profitieren. Wenn die Rezeptorexpression dagegen verloren geht, erhalten die Patienten ihre jetzt unwirksame zielgerichtete Therapie, bezahlen unnötige Kosten für die Medikation und sind den medikamentenassoziierten Nebenwirkungen ausgesetzt. Zum Beispiel kann die Einnahme eines Östrogen-Rezeptor-Modulators wie Tamoxifen das Risiko eines Endometriumkarzinoms in bestimmten Patientengruppen erhöhen (Early Breast Cancer Trialists' Collaborative, Darby et al. 2011).

Bislang gibt es jedoch keine randomisierten, kontrollierten, studienbasierten Belege für ein verbessertes Überleben von Patienten mit Therapieanpassungen entsprechend dem Rezeptorstatus ihrer Metastasen (Cardoso, Paluch-Shimon et al. 2020). Dennoch wird in den jüngsten Leitlinien empfohlen, (a) den Rezeptorstatus der metastasierten Läsion mindestens einmal zu bewerten und erneut zu testen (Carlson, Allred et al. 2012, Gennari, Andre et al. 2021) und (b) die Therapieschemata auf die Rezeptorexpression der metastasierten Läsion abzustimmen (Gennari, Andre et al. 2021) oder jede positive Rezeptorexpression aus Metastasen oder Primärtumorbiopsien zu berücksichtigen (Cardoso, Paluch-Shimon et al. 2020). Biopsien und Rebiopsien können jedoch nur dann durchgeführt werden, wenn die betreffende Läsion zugänglich ist und unerwünschte Wirkungen in einem angemessenen Verhältnis zu den potenziellen Vorteilen einer Therapieanpassung zu erwarten sind. Insbesondere bei Patienten mit Hirnmetastasen sind diese beiden Voraussetzungen oft schwer zu beurteilen und nur selten erfüllt. Die hier vorgeschlagene Methode zur nicht-invasiven Bestimmung des Rezeptorstatus auf der Grundlage von bildgebenden Biomarkern, die aus Standard-MR-Hirnschans abgeleitet werden, könnte daher (a) das therapeutische Management verbessern, indem sie eine Überwachung des Rezeptorstatus bei hohen Frequenzen ermöglicht und eine flexible Therapieoptimierung erlaubt, (b) das allgemeine Verständnis der Mechanismen der Rezeptorkonversion verbessern und (c) weitere Erkenntnisse über die Auswirkungen der Rezeptordivergenz und der angewandten zielgerichteten Therapieschemata auf das

Gesamtüberleben liefern. So konnten bestimmte Treibermutationen von Tumoren und pathologische Muster mit bestimmten radiomischen Signaturen in Verbindung gebracht werden (Lambin, Leijenaar et al. 2017, Limkin, Sun et al. 2017). Becker et al. konnten bei Patienten mit Lebermetastasen mithilfe von Radiomics Rückschlüsse auf die Prognose ziehen, wobei Metastasen mit höherer Entropie und geringerer Homogenität eine verbesserte Überlebenschance hatten (Beckers, Trebeschi et al. 2018).

In unserer Studie zeigten die Ergebnisse der Trainings-/Validierungsanalyse eine Überlegenheit des ER-Klassifikators (ROC AUC 0,82) im Vergleich zu den PR und HER2 -Klassifikatoren (ROC AUC 0,73 und 0,74). Die schlechteren Ergebnisse bei der Vorhersage des HER2-Status stehen im Einklang mit den in anderen Studien berichteten Leistungen bei der Klassifizierung von Primärtumoren (Luo, Xie et al. 2022), was auf das Vorhandensein von mehr heterogenen Bildgebungsmerkmalen bezogen auf den HER2-Status in diesen Läsionen hindeutet. Diese Hypothese wird durch den Vergleich der Werteverteilungen der 10 wichtigsten Histogramm-basierten Prädiktoren, die vom ER-Klassifikator verwendet werden, mit den Werteverteilungen von PR- und HER2-Metastasen gestützt: Entsprechende Prädiktoren in PR-positiven und HER2-positiven Läsionen weisen eine wesentlich höhere Variabilität und damit eine geringere Aussagekraft auf als in ER-positiven Tumoren. Zudem scheinen FLAIR-Sequenzen eher für die Differenzierung von ER-Metastasen geeignet zu sein; ER-positive Metastasen könnten ein spezifisches ödem-assoziiertes Muster bilden, welches für den maschinellen Lernalgorithmus sichtbar ist.

Bislang wurden nur wenige Ansätze für eine nicht-invasive Diagnostik der Rezeptorexpressionen von Metastasen untersucht. Eine Analyse in Bezug auf den Rezeptorstatus von Brustkrebsmetastasen im Gehirn wurde von Luo et al. (2022) durchgeführt (Luo, Xie et al. 2022): Es wurde ein logistisches Regressionsmodell trainiert, um den Rezeptorstatus von Hirnmetastasen vorherzusagen. Die Vorhersagen wurden mit Pathologieberichten der Metastasen und Primärtumoren verglichen. Die Studie ergab AUCs von 0,89, 0,88 und 0,87 für die ER, PR und HER2 -Status sowie eine Rezeptordiskordanz zwischen Metastase und Primarius von 27,5 % für ER/PR und 5 % für HER2. Es wurden ein Trainings- und ein Testdatensatz verwendet, aber die Kohorten waren relativ klein. Eine weitere Studie analysierte, ob eine molekulare Subtypenvorhersage anhand der radiomischen Signatur von primärem Brustkrebs möglich ist (Gong, Li et al. 2023). Hierfür wurden konventionelle und kontrastverstärkte Ultraschallaufnahmen verwendet. Obwohl mitunter recht gute Vorhersagegenauigkeiten erzielt worden sind (86,7% für triple negativ), wurden 6 Kategorien bei einer recht kleinen Probenanzahl unterschieden, was auf eine geringe Verallgemeinerbarkeit der Ergebnisse hindeutet. Zudem gilt die ultraschallbasierte Diagnostik als stark untersucherabhängig. Ein anderer Ansatz von Cao et al. (2022) untersuchte, ob habitat-basierte Radiomics in der Lage ist, den EGFR-Mutationsstatus in Hirnmetastasen von primären Lungenadenokarzinomen vorherzusagen, wobei das beste externe Validierungsmodell einen AUC von 0,9 erreichte (Cao, Pang et al. 2022). Positiv hervorzuheben bei dieser Studie ist, dass es sich um eine Multi-Center Studie handelte und eine recht große Patientenzahl eingeschlossen wurde.

Obwohl zahlreiche Assoziationen zwischen quantitativen Bildmerkmalen und molekularen Gewebemerkmale nachgewiesen wurden (Zhou, Scott et al. 2018), werden Radiomics-basierte künstliche Intelligenzsysteme immer noch häufig wegen potenziell nicht reproduzierbarer und unverständlicher Entscheidungswege kritisiert. Neben diesen allgemeinen Limitationen in radiomischen Studien weist unsere Untersuchung folgende spezifische Limitationen auf: (a) Wir haben die Rezeptorexpressionen nicht nach den Subtypen Luminal A, Luminal B, HER2-Überexpression und dem basalen Typ differenziert (Coates, Winer et al. 2015, Lev 2020). Eine solche Klassifizierung würde die Einbeziehung der Ki67-Vorhersage und ein wesentlich komplexeres Klassifizierungsdesign erfordern, um dem erforderlichen Multiklassenansatz Rechnung zu tragen. Da die Anzahl der Klassen und der Umfang der verfügbaren Trainingsstichproben maßgeblich die Vorhersageleistung bestimmen, würde ein solcher Ansatz entweder eine erhebliche Ausweitung des Stichprobenumfangs erfordern oder die Leistung des Klassifikators potenziell negativ beeinträchtigen. (b) Es kann nicht ausgeschlossen werden, dass einige Patienten in dem Kollektiv bereits während der Entwicklung der Hirnmetastasen einer systemischen Therapie zugeführt waren. In diesen Fällen war die Therapie nicht metastasenspezifisch wie zum Beispiel die Radiochirurgie, die hypothetisch zu einer Veränderung der radiomischen Signaturen führen könnte. (c) Wir definierten die Grundwahrheit für das Training der Klassifikatoren anhand von Pathologieberichten von Hirnmetastasen, sofern diese verfügbar waren, andernfalls verwendeten wir Primärtumorbiopsien. Dieses Vorgehen ist darauf zurückzuführen, dass metastasenspezifische Pathologieberichte nur für 38 der 106 eingeschlossenen Patienten verfügbar waren. Jedoch ist weder zwangsläufig davon auszugehen, dass der Rezeptorstatus zwischen Primarius und Metastase gleich ist, noch ist zwangsläufig davon auszugehen, dass der Rezeptorstatus zwischen verschiedenen Metastasen bei einem Patienten gleich ist (Hoefnagel, van der Groep et al. 2013, Schrijver, Suijkerbuijk et al. 2018). Geht man von einer Gesamt-Rezeptor-Konversionsrate von 20 % aus, so könnte eine entsprechend falsch definierte Grundwahrheit in 20 % der Fälle die beobachteten Klassifizierungsfehler von 16 % (ER+), 26 % (PR+) und 24 % (HER2+) teilweise erklären. (d) Niedrig aufgelöste Bilder wurden von der Studie ausgeschlossen, um die Aussagekraft der Texturanalyse zu verbessern. Die Einbeziehung niedrig aufgelöster Bilder würde die Vorhersageleistung potenziell verringern, da Texturmerkmale wesentlich zur Vorhersagekraft des Modells beitragen. (e) Eine Ausweitung der Anzahl eingeschlossener Patienten würde die Verallgemeinerbarkeit der Ergebnisse weiter erhöhen. Kleine Stichprobengrößen sind ein allgemeines Problem bei der Radiomics-Analyse und sind auf die begrenzte Verfügbarkeit von standardisierten und kommentierten multizentrischen Bildgebungsdatensätzen zurückzuführen. Dennoch deutet die durchgeführte Bewertung der Modellstabilität auf eine ausreichende Robustheit der Ergebnisse hin, um die Machbarkeit und die Grenzen Bildgebungsmarker-basierter Vorhersagen des Rezeptorstatus zu bewerten. In Zukunft könnten standardisierte Datenbanken wie The Cancer Imaging Archive (TCIA) oder Open-i (NLM) weiter ausgebaut und verbessert werden und so die Integration von bildgebenden Technologien, Radiomics und künstlicher Intelligenz in die klinische Routine erleichtern. Sie bieten eine solide Grundlage für die

Entwicklung und Validierung neuer diagnostischer Ansätze, fördern die Zusammenarbeit zwischen Forschern und verbessern letztendlich die klinische Versorgung von Krebspatienten.

Zukünftige Studien könnten das volle Potenzial der auf künstlicher Intelligenz basierenden Vorhersage des Rezeptorstatus durch eine wesentliche Vergrößerung der Studienpopulation, die Einbeziehung weiterer klinischer Befunde, die Verwendung standardisierter hochauflösender Bilder, die umfassende Integration ausschließlich metastasenbezogener histopathologischer Befunde und die Anpassung des Klassifikator-Designs für die Unterscheidung von Subtypen gemäß dem St. Gallen-Konsens untersuchen. Darüber hinaus könnten zukünftige Studien weitere Anwendungsbereiche erforschen, einschließlich der Vorhersage des Rezeptorstatus für Leber- und Knochenmetastasen und der Anpassung des Algorithmus an andere Primärtumore.

6 Schlussfolgerung

Wir stellten die Hypothese auf, dass quantitative Bildmerkmale, die aus multiparametrischen Standard-MR-Scans extrahiert wurden, zur nicht-invasiven Bestimmung des Rezeptorstatus von Hirnmetastasen bei Patienten und Patientinnen mit metastasiertem Brustkrebs verwendet werden können. Unsere Ergebnisse deuten darauf hin, dass auf maschinellem Lernen basierende Algorithmen den Rezeptorstatus mit AUCs von 0,82 für ER+ -Metastasen, 0,73 für PR+ -Metastasen und 0,74 für HER2+ -Metastasen vorhersagen können. Die beobachteten engen Konfidenzintervalle und geringen Standardabweichungen der ROC-Kurven trotz der Verwendung heterogener MRT-Daten (Scanner-Modell, Scanner-Protokoll, Auflösung) weisen auf eine hohe Stabilität der Vorhersageleistung in den Trainings-/Validierungssätzen hin. Vielversprechende AUC-Werte von deutlich über 0,5 in der unabhängigen Testgruppe bestätigen die ursprüngliche Hypothese. Unsere Studie zeigt, dass der vorgeschlagene Ansatz ein nicht-invasives Tracking der Rezeptorexpressionen bei Patienten mit metastasiertem Brustkrebs ermöglichen könnte.

Zukünftige Studien könnten das volle Potenzial der auf künstlicher Intelligenz basierenden Vorhersage des Rezeptorstatus bei Hirnmetastasen von Brustkrebspatientinnen erforschen und diesen Ansatz auf andere Bereiche wie Leber- oder Knochenmetastasen ausdehnen, was insgesamt die Aussichten für Patienten mit metastasierten Krebserkrankungen verbessern könnte.

7 Zusammenfassung

Purpose: Es wird die Machbarkeit der Unterscheidung von histologischen Subtypen basierend auf dem Rezeptorstatus (ER, PR, HER) von Brustkrebs-Hirnmastasen mithilfe von Radiomics MRT-Bildmerkmalen in binären Machine-Learning-Modellen für Patienten mit unbekanntem Rezeptorstatus zum Zeitpunkt der Diagnosestellung untersucht.

Materialien und Methoden: Die retrospektive Analyse basiert auf Bilddaten von 412 Hirnmetastasen von 106 Patienten, die zwischen September 2007 und September 2021 gewonnen wurden. Einschlusskriterien waren: diagnostizierte zerebrale Metastasen von Brustkrebs; histopathologische Berichte über den Progesteron- (PR), Östrogen- (ER) und humanen epidermalen Wachstumsfaktor 2- (HER2) Rezeptorstatus; und Verfügbarkeit von MR-Bildgebungsdaten. Insgesamt wurden 3367 quantitative Merkmale von T1-kontrastverstärkten, T1-nativen und FLAIR-Bildern und das entsprechende Patientenalter mit Hilfe von Random-Forest-Algorithmen bewertet. Als Bewertungsgrundlage wurde die Receiver-Operating-Characteristic Area-under-the-Curve (ROC AUC) verwendet. Die Bedeutung der einzelnen Merkmale wurde anhand von Gini-Impuritätsmaßen bewertet. Die Vorhersageleistung wurde mit 10 permutierten 5-fachen Kreuzvalidierungssätzen getestet, wobei die 30 wichtigsten Merkmale jedes Trainingssatzes verwendet wurden.

Ergebnisse: Die Receiver-Operating-Characteristic Areas-under-the-Curve der Validierungssätze betragen 0,82 (95% Konfidenzintervall [0,78; 0,85]) für ER+, 0,73 [0,69; 0,77] für PR+ und 0,74 [0,70; 0,78] für HER2+ und unterscheiden sich signifikant von einer Zufallsverteilung.

Schlussfolgerung: Durch den Einsatz von Machine-Learning-Techniken und der Analyse quantitativer Bildmerkmale aus Routine cMRT-Bildern ist es möglich, den Tumorrezeptorstatus von Hirnmetastasen bei Brustkrebspatienten mit hoher Genauigkeit vorherzusagen.

8 Abstract

Purpose: The feasibility of distinguishing histologic subtypes based on receptor status (ER, PR, HER) of breast cancer brain metastases using radiomics MR image features in binary machine learning models for patients with unknown receptor status at diagnosis is investigated.

Materials and Methods: This retrospective analysis is based on 412 brain metastases samples from 106 patients acquired between 09/2007 and 09/2021. Inclusion criteria were as follows: diagnosed cerebral metastases from breast cancer; histopathology reports on progesterone (PR), estrogen (ER), and human epidermal growth factor 2 (HER2) receptor status; and availability of MR imaging data. In total, 3367 quantitative features of T1 contrast-enhanced, T1 non-enhanced, and FLAIR images and corresponding patient age were evaluated utilizing random forest algorithms. The receiver operating characteristic area under the curve (ROC AUC) was used as the basis for evaluation. Feature importance was assessed using Gini impurity measures. Predictive performance was tested using 10 permuted 5-fold cross-validation sets employing the 30 most important features of each training set.

Results: Receiver operating characteristic areas under the curve of the validation sets were 0.82 (95% confidence interval [0.78; 0.85]) for ER+, 0.73 [0.69; 0.77] for PR+, and 0.74 [0.70; 0.78] for HER2+ and differ significantly from a random distribution.

Conclusion: Observations indicate that MR image features employed in a machine learning classifier could provide high discriminatory accuracy in predicting the receptor status of brain metastases from breast cancer.

9 Abkürzungsverzeichnis

CI	Konfidenzintervall
ER	Östrogenrezeptor
FLAIR	Fluid attenuated inversion recovery
GTV	Gross tumor volume
HER2	Humaner epidermaler Wachstumsfaktorrezeptor 2
MCC	Matthew Korrelationskoeffizient
ML	Machine Learning
NPV	Negativer prädiktiver Wert
PPV	Positiver prädiktiver Wert
PR	Progesteronrezeptor
ROC AUC	Receiver-Operating-Characteristic Area-under-the-Curve
ROI	Region of interest
T1 w CE	T1 kontrastverstärkt
T1w KM	T1 kontrastverstärkt
T1 w nativ	T1 nicht kontrastverstärkt
T1 w NE	T1 nicht kontrastverstärkt
VEGF	Vaskulärer endothelialer Wachstumsfaktor

10 Literaturverzeichnis

- A. Ramezan, C., T. A. Warner and A. E. Maxwell (2019). "Evaluation of Sampling and Cross-Validation Tuning Strategies for Regional-Scale Machine Learning Classification." Remote Sensing **11**(2): 185.
- Abbasian Ardakani, A., N. J. Bureau, E. J. Ciaccio and U. R. Acharya (2022). "Interpretation of radiomics features-A pictorial review." Comput Methods Programs Biomed **215**: 106609.
- Abraham, A., F. Pedregosa, M. Eickenberg, P. Gervais, A. Mueller, J. Kossaifi, A. Gramfort, B. Thirion and G. Varoquaux (2014). "Machine learning for neuroimaging with scikit-learn." Front Neuroinform **8**: 14.
- Aerts, H. J. (2016). "The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review." JAMA Oncol **2**(12): 1636-1642.
- Aurilio, G., L. Monfardini, S. Rizzo, A. Sciandivasci, L. Preda, V. Bagnardi, D. Disalvatore, G. Pruneri, E. Munzone, P. Della Vigna, G. Renne, M. Bellomi, G. Curigliano, A. Goldhirsch and F. Nole (2013). "Discordant hormone receptor and human epidermal growth factor receptor 2 status in bone metastases compared to primary breast cancer." Acta Oncol **52**(8): 1649-1656.
- Basler, L., H. S. Gabryś, S. A. Hogan, M. Pavic, M. Bogowicz, D. Vuong, S. Tanadini-Lang, R. Förster, K. Kudura, M. W. Huellner, R. Dummer, M. Guckenberger and M. P. Levesque (2020). "Radiomics, Tumor Volume, and Blood Biomarkers for Early Prediction of Pseudoprogression in Patients with Metastatic Melanoma Treated with Immune Checkpoint Inhibition." Clin Cancer Res **26**(16): 4414-4425.
- Beckers, R. C. J., S. Trebeschi, M. Maas, R. S. Schnerr, J. M. L. Sijmons, G. L. Beets, J. B. Houwers, R. G. H. Beets-Tan and D. M. J. Lambregts (2018). "CT texture analysis in colorectal liver metastases and the surrounding liver parenchyma and its potential as an imaging biomarker of disease aggressiveness, response and survival." Eur J Radiol **102**: 15-21.
- Bernard, S., L. Heutte and S. Adam (2009). Influence of Hyperparameters on Random Forest Accuracy. Multiple Classifier Systems, Berlin, Heidelberg, Springer Berlin Heidelberg.
- Bhatia, A., M. Birger, H. Veeraraghavan, H. Um, F. Tixier, A. S. McKenney, M. Cugliari, A. Caviasco, A. Bialczak, R. Malani, J. Flynn, Z. Zhang, T. J. Yang, B. D. Santomasso, A. N. Shoushtari and R. J. Young (2019). "MRI radiomic features are associated with survival in melanoma brain metastases treated with immune checkpoint inhibitors." Neuro Oncol **21**(12): 1578-1586.
- Breiman, L. (2001). "Random forests." Machine Learning **45**(1): 5-32.
- Cao, R., Z. Pang, X. Wang, Z. Du, H. Chen, J. Liu, Z. Yue, H. Wang, Y. Luo and X. Jiang (2022). "Radiomics evaluates the EGFR mutation status from the brain metastasis: a multi-center study." Phys Med Biol **67**(12).
- Cardoso, F., S. Paluch-Shimon, E. Senkus, G. Curigliano, M. S. Aapro, F. Andre, C. H. Barrios, J. Bergh, G. S. Bhattacharyya, L. Biganzoli, F. Boyle, M. J. Cardoso, L. A. Carey, J. Cortes, N. S. El Saghir, M. Elzayat, A. Eniu, L. Fallowfield, P. A. Francis, K. Gelmon, J. Gligorov, R. Haidinger, N. Harbeck, X. Hu, B. Kaufman, R. Kaur, B. E. Kiely, S. B. Kim, N. U. Lin, S. A. Mertz, S. Neciosup, B. V. Offersen, S. Ohno, O. Pagani, A. Prat, F. Penault-Llorca, H. S. Rugo, G. W. Sledge, C. Thomssen, D. A. Vorobiof, T. Wiseman, B. Xu, L. Norton, A. Costa and E. P. Winer (2020). "5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5)." Ann Oncol **31**(12): 1623-1649.
- Carlson, R. W., D. C. Allred, B. O. Anderson, H. J. Burstein, S. B. Edge, W. B. Farrar, A. Forero, S. H. Giordano, L. J. Goldstein, W. J. Gradishar, D. F. Hayes, C. A. Hudis, S. J. Isakoff, B. M. Ljung, D. A. Mankoff, P. K. Marcom, I. A. Mayer, B. McCormick, L. J. Pierce, E. C. Reed, M. L. Smith, H. Soliman, G. Somlo, R. L. Theriault, J. H. Ward, A. C. Wolff, R. Zellars, R. Kumar, D. A. Shead and N. National Comprehensive Cancer (2012).

"Metastatic breast cancer, version 1.2012: featured updates to the NCCN guidelines." *J Natl Compr Canc Netw* **10**(7): 821-829.

Chamberlain, M. C., C. S. Baik, V. K. Gadi, S. Bhatia and L. Q. Chow (2017). "Systemic therapy of brain metastases: non-small cell lung cancer, breast cancer, and melanoma." *Neuro Oncol* **19**(1): i1-i24.

Chicco, D. and G. Jurman (2023). "The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification." *BioData Mining* **16**(1): 4.

Coates, A. S., E. P. Winer, A. Goldhirsch, R. D. Gelber, M. Gnant, M. Piccart-Gebhart, B. Thürlimann, H.-J. Senn, P. Members and F. André (2015). "Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015." *Annals of oncology* **26**(8): 1533-1546.

Disha, R. A. and S. Waheed (2022). "Performance analysis of machine learning models for intrusion detection system using Gini Impurity-based Weighted Random Forest (GIWRF) feature selection technique." *Cybersecurity* **5**(1): 1.

Early Breast Cancer Trialists' Collaborative, G., S. Darby, P. McGale, C. Correa, C. Taylor, R. Arriagada, M. Clarke, D. Cutter, C. Davies, M. Ewertz, J. Godwin, R. Gray, L. Pierce, T. Whelan, Y. Wang and R. Peto (2011). "Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials." *Lancet* **378**(9804): 1707-1716.

Fonov, V., A. C. Evans, K. Botteron, C. R. Almlri, R. C. McKinstry, D. L. Collins and G. Brain Development Cooperative (2011). "Unbiased average age-appropriate atlases for pediatric studies." *Neuroimage* **54**(1): 313-327.

Galavis, P. E., C. Hollensen, N. Jallow, B. Paliwal and R. Jeraj (2010). "Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters." *Acta Oncol* **49**(7): 1012-1016.

Gennari, A., F. Andre, C. H. Barrios, J. Cortes, E. de Azambuja, A. DeMichele, R. Dent, D. Fenlon, J. Gligorov, S. A. Hurvitz, S. A. Im, D. Krug, W. G. Kunz, S. Loi, F. Penault-Llorca, J. Ricke, M. Robson, H. S. Rugo, C. Saura, P. Schmid, C. F. Singer, T. Spanic, S. M. Tolaney, N. C. Turner, G. Curigliano, S. Loibl, S. Paluch-Shimon, N. Harbeck and E. G. C. E. a. clinicalguidelines@esmo.org (2021). "ESMO Clinical Practice Guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer." *Ann Oncol* **32**(12): 1475-1495.

Gennari, A., F. Andre, C. H. Barrios, J. Cortes, E. de Azambuja, A. DeMichele, R. Dent, D. Fenlon, J. Gligorov, S. A. Hurvitz, S. A. Im, D. Krug, W. G. Kunz, S. Loi, F. Penault-Llorca, J. Ricke, M. Robson, H. S. Rugo, C. Saura, P. Schmid, C. F. Singer, T. Spanic, S. M. Tolaney, N. C. Turner, G. Curigliano, S. Loibl, S. Paluch-Shimon, N. Harbeck and E. G. Comm (2021). "ESMO Clinical Practice Guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer." *Annals of Oncology* **32**(12): 1475-1495.

Gillies, R. J., P. E. Kinahan and H. Hricak (2016). "Radiomics: Images Are More than Pictures, They Are Data." *Radiology* **278**(2): 563-577.

Gong, X., Q. Li, L. Gu, C. Chen, X. Liu, X. Zhang, B. Wang, C. Sun, D. Yang, L. Li and Y. Wang (2023). "Conventional ultrasound and contrast-enhanced ultrasound radiomics in breast cancer and molecular subtype diagnosis." *Front Oncol* **13**: 1158736.

Guiot, J., A. Vaidyanathan, L. Deprez, F. Zerka, D. Danthine, A. N. Frix, P. Lambin, F. Bottari, N. Tsoutzidis, B. Miraglio, S. Walsh, W. Vos, R. Hustinx, M. Ferreira, P. Lovinfosse and R. T. H. Leijenaar (2022). "A review in radiomics: Making personalized medicine a reality via routine imaging." *Med Res Rev* **42**(1): 426-440.

Haralick, R., K. Shanmugam and I. Dinstein (1973). "Textural Features for Image Classification." *IEEE Trans Syst Man Cybern* **SMC-3**: 610-621.

Hoefnagel, L. D., P. van der Groep, M. J. van de Vijver, J. E. Boers, P. Wesseling, J. Wesseling, C. Dutch Distant Breast Cancer Metastases, E. van der Wall and P. J. van Diest (2013). "Discordance in ERalpha, PR and HER2 receptor status across different distant breast cancer metastases within the same patient." *Ann Oncol* **24**(12): 3017-3023.

Hosonaga, M., H. Saya and Y. Arima (2020). "Molecular and cellular mechanisms underlying brain metastasis of breast cancer." *Cancer Metastasis Rev* **39**(3): 711-720.

Hu, L. S. and K. R. Swanson (2020). "Roadmap for the clinical integration of radiomics in neuro-oncology." *Neuro Oncol* **22**(6): 743-745.

Hu, L. S., H. Yoon, J. M. Eschbacher, L. C. Baxter, A. C. Dueck, A. Nespodzany, K. A. Smith, P. Nakaji, Y. Xu, L. Wang, J. P. Karis, A. J. Hawkins-Daarud, K. W. Singleton, P. R. Jackson, B. J. Anderies, B. R. Bendok, R. S. Zimmerman, C. Quarles, A. B. Porter-Umphrey, M. M. Mrugala, A. Sharma, J. M. Hoxworth, M. G. Sattur, N. Sanai, P. E. Koulemberis, C. Krishna, J. R. Mitchell, T. Wu, N. L. Tran, K. R. Swanson and J. Li (2019). "Accurate Patient-Specific Machine Learning Models of Glioblastoma Invasion Using Transfer Learning." *AJNR Am J Neuroradiol* **40**(3): 418-425.

Iglesias, J. E., C. Y. Liu, P. M. Thompson and Z. Tu (2011). "Robust brain extraction across datasets and comparison with publicly available methods." *IEEE Trans Med Imaging* **30**(9): 1617-1634.

Jung, J., S. H. Lee, M. Park, J. H. Youn, S. H. Shin, H. S. Gwak and H. Yoo (2018). "Discordances in ER, PR, and HER2 between primary breast cancer and brain metastasis." *Journal of Neuro-Oncology* **137**(2): 295-302.

Kanchan, R. K., J. A. Siddiqui, S. Mahapatra, S. K. Batra and M. W. Nasser (2020). "microRNAs Orchestrate Pathophysiology of Breast Cancer Brain Metastasis: Advances in Therapy." *Mol Cancer* **19**(1): 29.

Karpievitch, Y. V., E. G. Hill, A. P. Leclerc, A. R. Dabney and J. S. Almeida (2009). "An introspective comparison of random forest-based classifiers for the analysis of cluster-correlated data by way of RF++." *PLoS One* **4**(9): e7087.

Kniep, H. C. (2020). *Radiomics of Brain MRI: Utility in Prediction of Metastatic Tumor Type* Universitätsklinikum Hamburg-Eppendorf.

Kniep, H. C., F. Madesta, T. Schneider, U. Hanning, M. H. Schonfeld, G. Schon, J. Fiehler, T. Gauer, R. Werner and S. Gellissen (2019). "Radiomics of Brain MRI: Utility in Prediction of Metastatic Tumor Type." *Radiology* **290**(2): 479-487.

Kroigard, A. B., M. J. Larsen, M. Thomassen and T. A. Kruse (2016). "Molecular Concordance Between Primary Breast Cancer and Matched Metastases." *Breast J* **22**(4): 420-430.

Kumar, V., Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. Aerts, A. Dekker, D. Fenstermacher, D. B. Goldgof, L. O. Hall, P. Lambin, Y. Balagurunathan, R. A. Gatenby and R. J. Gillies (2012). "Radiomics: the process and the challenges." *Magn Reson Imaging* **30**(9): 1234-1248.

Lambin, P., R. T. H. Leijenaar, T. M. Deist, J. Peerlings, E. E. C. de Jong, J. van Timmeren, S. Sanduleanu, R. T. H. M. Larue, A. J. G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F. M. Mottaghy, J. E. Wildberger and S. Walsh (2017). "Radiomics: the bridge between medical imaging and personalized medicine." *Nature Reviews Clinical Oncology* **14**(12): 749-762.

Lev, S. (2020). "Targeted therapy and drug resistance in triple-negative breast cancer: the EGFR axis." *Biochem Soc Trans* **48**(2): 657-665.

Limkin, E. J., R. Sun, L. Dercle, E. I. Zacharaki, C. Robert, S. Reuzé, A. Schernberg, N. Paragios, E. Deutsch and C. Ferte (2017). "Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology." *Ann Oncol* **28**(6): 1191-1206.

Lo Gullo, R., I. Daimiel, E. A. Morris and K. Pinker (2020). "Combining molecular and imaging metrics in cancer: radiogenomics." *Insights into Imaging* **11**(1): 1.

Louppe, G. (2013). "Understanding variable importances in forests of randomized trees." Advances in Neural Information Processing Systems

Luo, X., H. Xie, Y. Yang, C. Zhang, Y. Zhang, Y. Li, Q. Yang, D. Wang, Y. Luo, Z. Mai, C. Xie and S. Yin (2022). "Radiomic Signatures for Predicting Receptor Status in Breast Cancer Brain Metastases." Front Oncol **12**: 878388.

Mason, S. J. and N. E. Graham (2002). "Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation." Quarterly Journal of the Royal Meteorological Society **128**(584): 2145-2166.

Mayerhoefer, M. E., A. Materka, G. Langs, I. Haggstrom, P. Szczypinski, P. Gibbs and G. Cook (2020). "Introduction to Radiomics." J Nucl Med **61**(4): 488-495.

Michel, A., M. D. Oppong, L. Rauschenbach, D. Pierscianek, T. F. Dinger, T. Schmidt, J. Hense, C. Pottgen, R. Kimmig, Y. Ahmadipour, N. Ozkan, O. Muller, A. Junker, U. Sure, R. Jabbarli and N. El Hindy (2021). "HER2 Receptor Conversion Is a strong Survival Predictor in Patients with Breast Cancer Brain Metastases." World Neurosurg **152**: e332-e343.

Powers, D. M. (2020). "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." arXiv preprint arXiv:2010.16061.

Probst, P., B. Bischl and A.-L. Boulesteix (2018). Tunability: Importance of Hyperparameters of Machine Learning Algorithms.

Reinhold, J. C., B. E. Dewey, A. Carass and J. L. Prince (2019). "Evaluating the Impact of Intensity Normalization on MR Image Synthesis." Proc SPIE Int Soc Opt Eng **10949**.

Rios Velazquez, E., C. Parmar, Y. Liu, T. P. Coroller, G. Cruz, O. Stringfield, Z. Ye, M. Makrigiorgos, F. Fennessy, R. H. Mak, R. Gillies, J. Quackenbush and H. Aerts (2017). "Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer." Cancer Res **77**(14): 3922-3930.

Robin, X., N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J. C. Sanchez and M. Muller (2011). "pROC: an open-source package for R and S+ to analyze and compare ROC curves." BMC Bioinformatics **12**: 77.

Sarica, A., A. Cerasa and A. Quattrone (2017). "Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review." Front Aging Neurosci **9**: 329.

Scapicchio, C., M. Gabelloni, A. Barucci, D. Cioni, L. Saba and E. Neri (2021). "A deep look into radiomics." La radiologia medica **126**(10): 1296-1311.

Schrijver, W., P. Selenica, J. Y. Lee, C. K. Y. Ng, K. A. Burke, S. Piscuoglio, S. H. Berman, J. S. Reis-Filho, B. Weigelt, P. J. van Diest and C. B. Moelans (2018). "Mutation Profiling of Key Cancer Genes in Primary Breast Cancers and Their Distant Metastases." Cancer Res **78**(12): 3112-3121.

Schrijver, W., K. P. M. Suijkerbuijk, C. H. van Gils, E. van der Wall, C. B. Moelans and P. J. van Diest (2018). "Receptor Conversion in Distant Breast Cancer Metastases: A Systematic Review and Meta-analysis." J Natl Cancer Inst **110**(6): 568-580.

Senkus, E., S. Kyriakides, S. Ohno, F. Penault-Llorca, P. Poortmans, E. Rutgers, S. Zackrisson, F. Cardoso and E. G. Committee (2015). "Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up." Ann Oncol **26 Suppl 5**: v8-30.

Sidey-Gibbons, J. A. M. and C. J. Sidey-Gibbons (2019). "Machine learning in medicine: a practical introduction." BMC Medical Research Methodology **19**(1): 64.

Tustison, N. J., B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich and J. C. Gee (2010). "N4ITK: improved N3 bias correction." IEEE Trans Med Imaging **29**(6): 1310-1320.

Ute-Susann Albert, W. B., Markus Follmann, Wolfgang Janni, Ina Kopp (2021). Interdisziplinäre S3-Leitlinie für die Früherkennung, Diagnostik, Therapie und Nachsorge des Mammakarzinoms.

van Griethuysen, J. J. M., A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. H. Beets-Tan, J. C. Fillion-Robin, S. Pieper and H. Aerts (2017). "Computational Radiomics System to Decode the Radiographic Phenotype." Cancer Res **77**(21): e104-e107.

Wang, J. Z. (2001). "Wavelets and imaging informatics: A review of the literature." Journal of Biomedical Informatics **34**(2): 129.

Wang, Y., F. Ye, Y. Liang and Q. Yang (2021). "Breast cancer brain metastasis: insight into molecular mechanisms and therapeutic strategies." Br J Cancer **125**(8): 1056-1067.

Watase, C., S. Shiino, T. Shimoi, E. Noguchi, T. Kaneda, Y. Yamamoto, K. Yonemori, S. Takayama and A. Suto (2021). "Breast Cancer Brain Metastasis-Overview of Disease State, Treatment Options and Future Perspectives." Cancers (Basel) **13**(5).

Yang, M., M. E. Forbes, R. L. Bitting, S. S. O'Neill, P. C. Chou, U. Topaloglu, L. D. Miller, G. A. Hawkins, S. C. Grant, B. R. DeYoung, W. J. Petty, K. Chen, B. C. Pasche and W. Zhang (2018). "Incorporating blood-based liquid biopsy information into cancer staging: time for a TNMB system?" Ann Oncol **29**(2): 311-323.

Yang, S. X., N. Wu, L. Zhang and M. Li (2022). "[The influence of CT acquisition and reconstruction parameters on the stability of radiomic features of pure ground-glass nodules]." Zhonghua Zhong Liu Za Zhi **44**(9): 981-986.

Yi, Z. B., P. Yu, S. Zhang, W. N. Wang, Y. Q. Han, Q. C. Ouyang, M. Yan, X. J. Wang, X. C. Hu, Z. F. Jiang, T. Huang, Z. S. Tong, S. S. Wang, Y. M. Yin, H. Li, R. X. Yang, H. W. Yang, Y. E. Teng, T. Sun, L. Cai, H. Y. Li, X. Chen, J. J. He, X. L. Liu, S. E. Yang, J. Y. Wang, J. H. Fan, Y. L. Qiao and B. H. Xu (2021). "Profile and outcome of receptor conversion in breast cancer metastases: a nation-wide multicenter epidemiological study." Int J Cancer **148**(3): 692-701.

Zhang, Z., M. W. Beck, D. A. Winkler, B. Huang, W. Sibanda and H. Goyal (2018). "Opening the black box of neural networks: methods for interpreting neural network models in clinical applications." Ann Transl Med **6**(11): 216.

Zhao, W., L. Sun, G. Dong, X. Wang, Y. Jia and Z. Tong (2021). "Receptor conversion impacts outcomes of different molecular subtypes of primary breast cancer." Ther Adv Med Oncol **13**: 17588359211012982.

Zhou, M., J. Scott, B. Chaudhury, L. Hall, D. Goldgof, K. W. Yeom, M. Iv, Y. Ou, J. Kalpathy-Cramer, S. Napel, R. Gillies, O. Gevaert and R. Gatenby (2018). "Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches." AJNR Am J Neuroradiol **39**(2): 208-216.

11 Erklärung des Eigenanteils

Ich, Alexander Heitkamp, habe im Wesentlichen folgende Aufgaben bei der Erstellung der Dissertation übernommen beziehungsweise an folgenden Aufgaben mitgewirkt: Vervollständigung und Kuration der Bilddatenbank, Planung der Forschungsfrage, Anpassung der Python-Skripte hinsichtlich Bildprozessierung und Machine Learning, statistische Auswertung, Verfassen des Manuskripts, Korrespondenz mit dem peer-reviewed, open access Journal Cancers. Ich bin alleiniger Erstautor der vorliegenden Promotionsarbeit.

Im Folgenden wird die Konzeption der Promotionsarbeit im Detail skizziert: Ich habe die Patientendatenbank, die ich von Herrn Dr. Gauer und der Strahlentherapie erhalten habe, durch eigene Suchanfragen im PACS vervollständigt. Gemeinsam mit Frau PD Dr. Gellißen und Herrn Dr. Kniep haben wir die Forschungsfrage konzipiert. Gemäß der zuvor festgelegten Ein- und Ausschlusskriterien habe ich die Patientendatenbank selektiert. Unter vorheriger Anleitung habe ich die Hirnmetastasen mitsegmentiert. Meine Segmentierungen wurden durch Frau PD Dr. Gellißen und Herrn Dr. Kniep validiert. Weiterhin habe ich die Schritte der Bilddatenprozessierung (Registrierung, Gehirn-Extraktion, N4ITK-Bias-Korrektur, Intensitätsnormalisierung) durchgeführt. Die Bash- und Pythonskripte hierfür habe ich teilweise selbst erstellt und teilweise von Herrn Dr. Kniep erhalten und an unsere Daten angepasst. Die Merkmals-Extraktion der MRT-Bilder habe ich mit einem Python-Skript auf Basis von Pyradiomics durchgeführt. Einen Vorschlag hierfür habe ich von Herrn Dr. Kniep erhalten, diesen habe ich weiterentwickelt und an unsere Daten angepasst. Gleiches gilt für die Feature Selektion und für die Implementierung der Random Forest Algorithmen auf Basis der Python-Bibliothek Scikit-Learn. Die statistische Auswertung habe ich unter Supervision selbstständig mit R und Python durchgeführt. Beim Verfassen des Manuskripts wurde ich insbesondere von Herrn Dr. Kniep, Herrn Dr. Gauer und Herrn PD Dr. Werner unterstützt und umfassend beraten.

12 Danksagung

Mit großer Freude und Dankbarkeit möchte ich an dieser Stelle anlässlich meiner Promotion die Gelegenheit nutzen, mich bei all den Menschen zu bedanken, die mich auf meinem akademischen Weg unterstützt und inspiriert haben. Ohne ihre wertvolle Unterstützung und Ermutigung wäre diese Leistung nicht realisierbar gewesen.

Zunächst gilt mein großer Dank Herrn Dr. Helge Kniep, der mich von Themenfindung und medizinischen sowie informatischen Fragestellungen bis hin zu organisatorischen Herausforderungen fortwährend unterstützt hat. Ich danke ihm sowie Frau PD Dr. Susanne Gellißen für die Verfügungstellung eines anspruchsvollen, aber sehr interessanten Promotionsthemas. Darüber hinaus möchte ich mich für die Geduld, den Optimismus und die positive Arbeitsatmosphäre bei Herrn Dr. Helge Kniep bedanken, die mich auch bei komplexen Anforderungen durchgehend motiviert haben.

Weiterhin danke ich Herrn Prof. Jens Fiehler für die Unterstützung seiner Abteilung ohne dessen Rahmen die Erstellung dieser Arbeit nicht möglich gewesen wäre. Zudem möchte ich mich für die inhaltlichen Revisionen beim Manuskript bedanken.

Frau PD Dr. Susanne Gellißen möchte ich außerdem für die inhaltliche Hintergrundunterstützung, für die radiologische Expertise sowie für die Konzeption der Fragestellung in Zusammenarbeit mit Dr. Helge Kniep danken.


Ich danke Herrn Dr. Tobias Gauer und der Strahlentherapie für die Bereitstellung der Daten sowie für interessante Gespräche hinsichtlich weiterer Fragestellungen um das Forschungsthema. Ich danke ihm für die kritische Auseinandersetzung mit den Workflows dieser Promotion und für seine Hilfe beim Abfassen des Manuskripts.

Mein weiterer Dank gilt PD Dr. René Werner, der mir anfangs die nötige Zuversicht vermittelt hat, eine programmierbasierte Promotionsarbeit zu beginnen und mich weiterhin bei technischen Schwierigkeiten sowie beim Verfassen des Manuskripts unterstützt hat.

Ganz besonderer Dank gilt meinen Eltern, meinem Bruder, meinen Großeltern und meiner Freundin, die mich auf meinen Wegen und Entscheidungen immer begleiten und an mich glauben.

Insbesondere meinen Eltern danke ich für die fortwährende Unterstützung und das entgegengebrachte Vertrauen und dass sie mir diese Arbeit ermöglichen.

13 Lebenslauf



**ALEXANDER JOACHIM
GIESBERT HEITKAMP**

PHYSICIAN

PROFILE

Trained as a physician and industrial engineer with a special interest in radiology and life science, economic relations and machine learning.

COMPETENCES

- Teamwork
- Resilience
- Ability to Travel
- Analytical Thinking
- Quick Learner

SKILLS

CODING/ PYTHON

MICROSOFT OFFICE

STATISTICS R

alekzanderheitkamp@googlemail.com
Preystraße 1, 22303 Hamburg
+49 151 51137933
17. June 1992
German

WORK EXPERIENCE

- Sep 2018 - Sep 2023
WORKING STUDENT IN THE PEDIATRIC INTENSIVE CARE UNIT
Universitätsklinikum Hamburg-Eppendorf
Hamburg, Germany
Administration, procurement of drugs and medical devices, shift scheduling
- May 2022 - May 2023
PRACTICAL YEAR
Medical School
Hamburg, Germany; Zurich, Switzerland
Radiology, Cardiology, Surgery, Stem Cell Transplantation and Neurosurgery
- Jun 2016 - Dec 2016
WORKING STUDENT
Otto Group GmbH
Hamburg, Germany
Corporate Finance, Treasury

ACADEMIC CAREER

- Oct 2021 - Sep 2023
DISSERTATION, NEURORADIOLOGY
Universitätsklinikum Hamburg-Eppendorf
Radiomics, Machine Learning, Oncology
- Oct 2016 - Jun 2023
APPROBATION, MEDICAL SCHOOL
University of Hamburg
First Study Section: 1.0
Second Study Section: 2.0
- Oct 2012 - Sep 2016
BACHELOR OF SCIENCE, INDUSTRIAL ENGINEERING
University of Hamburg
Major: Finance and Investing
- Aug 2002 - Jun 2011
A-LEVELS
Gymnasium Canisianum
Lüdinghausen, Germany

LANGUAGE SKILLS

- English | advanced
- Spanish | basic knowledge
- French | basic knowledge

INTERNATIONAL EXPERIENCE

- San Diego, USA | 5 month
- Barcelona, Spain | 3 month
- Zurich, Switzerland | 4 month

Hamburg,

14 Eidesstattliche Erklärung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: