

## Article

# Unsupervised Analysis Based on DCE-MRI Radiomics Features Revealed Three Novel Breast Cancer Subtypes with Distinct Clinical Outcomes and Biological Characteristics

Wenlong Ming<sup>1,2</sup>, Fuyu Li<sup>1</sup>, Yanhui Zhu<sup>3</sup>, Yunfei Bai<sup>1</sup> , Wanjun Gu<sup>1,4</sup>, Yun Liu<sup>5</sup>, Xiaoan Liu<sup>3</sup>, Xiao Sun<sup>1,\*</sup>  and Hongde Liu<sup>1,\*</sup> 

<sup>1</sup> State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China

<sup>2</sup> Division of Medical Image Computing, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

<sup>3</sup> Department of Breast Surgery, the First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, China

<sup>4</sup> Collaborative Innovation Center of Jiangsu Province of Cancer Prevention and Treatment of Chinese Medicine, School of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing 210023, China

<sup>5</sup> Department of Information, the First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, China

\* Correspondence: xsun@seu.edu.cn (X.S.); liuhongde@seu.edu.cn (H.L.)



**Citation:** Ming, W.; Li, F.; Zhu, Y.; Bai, Y.; Gu, W.; Liu, Y.; Liu, X.; Sun, X.; Liu, H. Unsupervised Analysis Based on DCE-MRI Radiomics Features Revealed Three Novel Breast Cancer Subtypes with Distinct Clinical Outcomes and Biological Characteristics. *Cancers* **2022**, *14*, 5507. <https://doi.org/10.3390/cancers14225507>

Academic Editors: Katja Pinker-Domenig and Antonis Valachis

Received: 27 September 2022

Accepted: 7 November 2022

Published: 9 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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/>).

**Simple Summary:** Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is an important approach for the diagnosis and evaluation of breast cancer (BC) in clinical practice. Recently, DCE-MRI-based radiomics studies have received widespread attention and application in BC research, such as in non-invasively predicting subtypes and recurrence risks. Therefore, we collected two radiogenomics cohorts of BC and identified and validated three novel imaging subtypes by unsupervised analysis in this work. In several external datasets, we found that breast tumors with larger sizes and showing rapid enhancement patterns generally had the worst prognostic outcomes. The bioinformatics analysis revealed significant differences in gene expression profiling and tumor microenvironment characteristics among the three imaging subtypes. These findings highlight the heterogeneity in BC imaging and its potential value as a clinical biomarker for BC and for achieving precision medicine in BC.

**Abstract:** Background: This study aimed to reveal the heterogeneity of dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) of breast cancer (BC) and identify its prognosis values and molecular characteristics. Methods: Two radiogenomics cohorts ( $n = 246$ ) were collected and tumor regions were segmented semi-automatically. A total of 174 radiomics features were extracted, and the imaging subtypes were identified and validated by unsupervised analysis. A gene-profile-based classifier was developed to predict the imaging subtypes. The prognostic differences and the biological and microenvironment characteristics of subtypes were uncovered by bioinformatics analysis. Results: Three imaging subtypes were identified and showed high reproducibility. The subtypes differed remarkably in tumor sizes and enhancement patterns, exhibiting significantly different disease-free survival (DFS) or overall survival (OS) in the discovery cohort ( $p = 0.024$ ) and prognosis datasets ( $p$  ranged from  $<0.0001$  to  $0.0071$ ). Large sizes and rapidly enhanced tumors usually had the worst outcomes. Associations were found between imaging subtypes and the established subtypes or clinical stages ( $p$  ranged from  $<0.001$  to  $0.011$ ). Imaging subtypes were distinct in cell cycle and extracellular matrix (ECM)-receptor interaction pathways (false discovery rate, FDR  $< 0.25$ ) and different in cellular fractions, such as cancer-associated fibroblasts ( $p < 0.05$ ). Conclusions: The imaging subtypes had different clinical outcomes and biological characteristics, which may serve as potential biomarkers.

**Keywords:** radiogenomics; radiomics; breast cancer; dynamic contrast-enhanced magnetic resonance imaging; tumor heterogeneity; prognosis

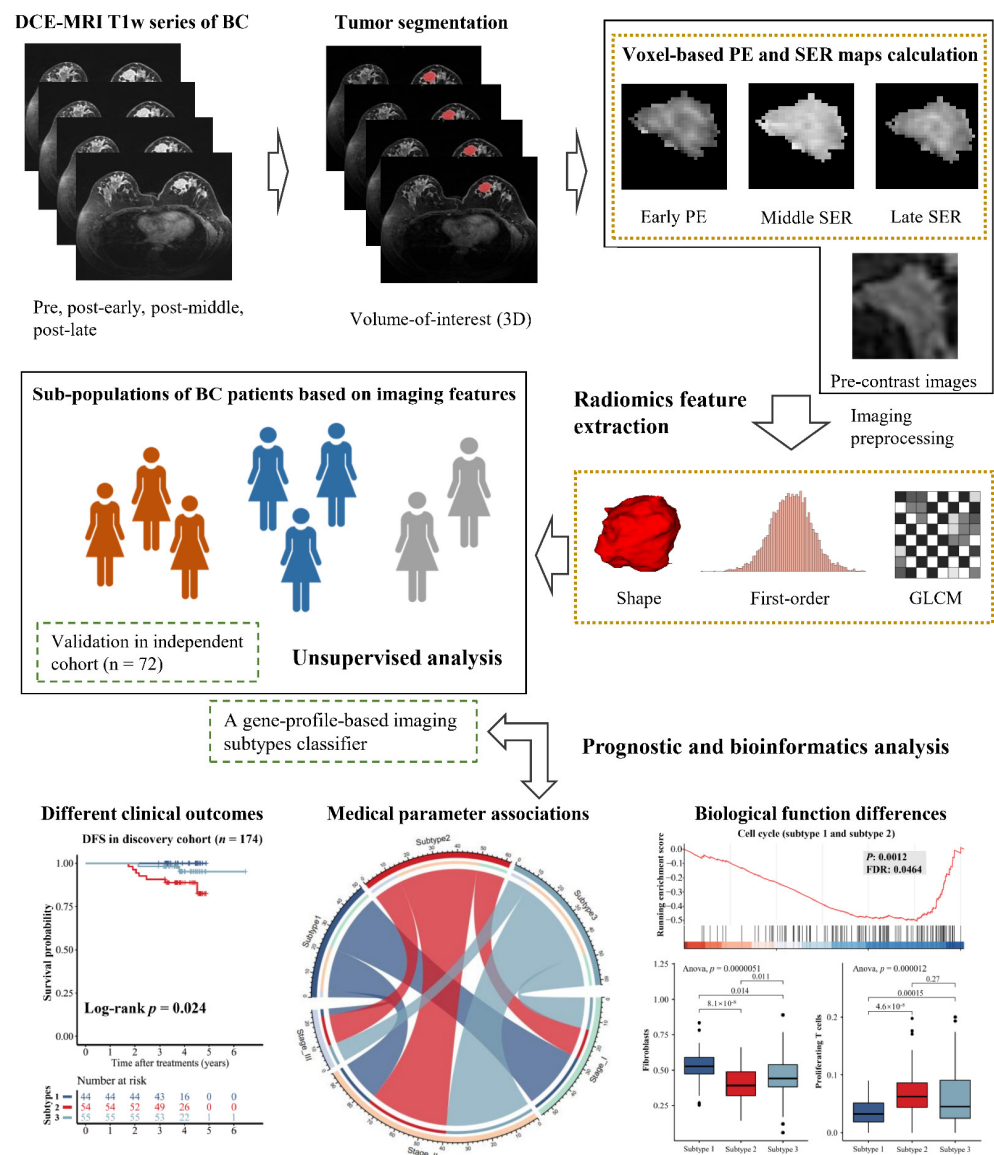
---

## 1. Introduction

Breast cancer (BC) has surpassed lung cancer as the most prevalent cancer in the world, placing a heavy burden on global healthcare each year [1]. The heterogeneity of BC has been noted in gene expression profiles, histology, and clinical outcomes for a long time, which have served as the basis of disease classification [2,3]. Based on the gene expression profile, BC can be defined as five PAM50 intrinsic molecular subtypes, including luminal-A, luminal-B, HER2-enriched, basal-like, and normal-like [4,5]. In the routine implementation, various histopathological receptors including estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki67 are widely used to determine the clinical subtypes of BC, and guide the clinical decision-making [6,7]. These factors shape the distinct patterns of tumor proliferation, metastasis, and treatment resistance in patients.

Many studies have been performed to dissect the molecular and clinical heterogeneity of BC, but that alone is still not enough in the era of personalized medicine. Fortunately, the development of multi-omics offers the possibility to understand the whole spectrum of disease, especially using radiomics and radiogenomics, providing an emerging perspective to characterize a disease noninvasively [8,9]. Radiomics is now gaining increasing attention in cancer research, by integrating high-throughput quantitative medical imaging features into clinical decision support systems, such as malignancy diagnosis, clinical parameters prediction, and prognosis and treatment response prediction [10–18]. For example, a radiomics signature consisting of 24 computed tomography imaging features performed well in predicting lymph node metastasis in colorectal cancer [18], and recent work showed that magnetic resonance imaging (MRI) features were associated with the genomic subclones and could predict the clinical outcomes of BC [19]. In the clinic, various imaging techniques are applied to diagnose BC, such as ultrasound, mammography, and MRI; among these, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) data are widely used in BC radiomics or radiogenomics studies for their advantages in high quality and three-dimension resolution. However, most studies focus on uncovering the relationship between medical imaging features and the molecular or clinical characteristics of the disease, and few have analyzed the disease heterogeneity in imaging phenotypes from an independent perspective [19–25].

Although imaging features are important phenotypes, few studies have focused on disease heterogeneity in radiomics compared to genomics or transcriptomics. In this study, we hypothesized that imaging features could reflect BC heterogeneity. Therefore, we recruited two independent radiomics cohorts with both DCE-MRI and RNA-Seq data in this work, one as the imaging-subtype discovery cohort with 174 samples and the other as the imaging-subtype validation cohort with 72 samples. The representative imaging features from tumor regions were extracted to mine the potential imaging subtypes *de novo*. We further identified differences in clinical examination, imaging, and prognosis among the subtypes, and revealed the underlying reasons for this heterogeneity in terms of transcriptional activity and tumor immune microenvironment. The workflow is shown in Figure 1. Our findings demonstrated that the imaging subtypes with distinct clinical and molecular characteristics were reliable and reproducible, and were useful for the noninvasive prediction of outcome and biological functions of BC in the clinic.



**Figure 1.** Workflow of this unsupervised analysis based on DCE-MRI radiomics features in BC patients. In this study, we took advantage of DCE-MRI techniques and calculated the voxel-based percentage enhancement (PE) and signal enhancement ratio (SER) maps for each patient. Two independent radiogenomics cohorts ( $n = 246$ ) were collected to identify and validate the novel imaging subtypes. The clinical and biological associations of DCE-MRI subtypes were further investigated.

## 2. Materials and Methods

### 2.1. Study Design

This study was designed as a multicentric exploratory study to investigate and validate the heterogeneity of quantitative DCE-MRI radiomics features in BC and to uncover the potential values in the prognosis of this kind of heterogeneity and the biological mechanisms behind it (Figure 1). To achieve this goal, an unsupervised analysis pipeline including the volume-of-interest (VOI) segmentation of tumor lesions, voxel-based percentage enhancement (PE) and signal enhancement ratio (SER) maps calculation, image normalization, resampling, radiomics feature extraction, and consensus clustering and bioinformatics analysis was conducted in two independent radiogenomics cohorts. The patients and technical details are reported below.

## 2.2. Radiogenomics Cohorts and Datasets

Two independent radiogenomics cohorts with both DCE-MRI and RNA-Seq data (as imaging-subtype discovery and validation cohorts, respectively) were recruited. Data from the imaging-subtype discovery cohort included preoperative T1-weighted DCE-MRI and RNA-Seq data collected between August 2016 and December 2018 in the local database. The imaging-subtype validation cohort is a public dataset collected from 4 medical centers, including the Memorial Sloan Kettering Cancer Center, the Mayo Clinic, the University of Pittsburgh Medical Center, and the Roswell Park Cancer Institute [26]. The DCE-MRI data were from the Cancer Imaging Archive (TCIA) and the RNA-seq data were part of a larger prognosis validation dataset, TCGA-BRCA ( $n = 1050$ ), from the Cancer Genome Atlas (TCGA). The inclusion and exclusion criteria are displayed in Figure S1. Finally, the imaging-subtype discovery and validation cohorts included 174 and 72 patients, respectively. In addition, six external datasets of BC with only gene expression profiles ( $n = 1443$ ) were downloaded from the Gene Expression Omnibus (GEO) database, including GSE1456, GSE3494, GSE7390, GSE20685, GSE25055, and GSE25065, which were used as prognosis validation datasets in this study.

## 2.3. Pathological Assessment

The clinical immunohistochemistry (IHC) results of ER, PR, HER2, Ki67 and IHC subtypes were determined for the patients in the discovery cohort. A patient with ER positive, HER2 negative, high PR expression (more than 20%), and low Ki67 expression (less than 20%) was regarded as luminal-A. Luminal-B patients were ER positive, HER2 negative, with low PR expression or high Ki-67 expression, or both ER and HER2 positive. ER, PR negative, and HER2 positive were HER2-positive, and, finally, patients with all negative IHC receptors were defined as triple negative BC (TNBC).

## 2.4. Imaging Parameters

The T1-weighted images were scanned in the axial position and acquired by Siemens TrioTim 3-Tesla scanner (Siemens Healthcare, Erlangen, Germany) in the imaging-subtype discovery cohort. The detailed parameters of most images were as follows: flip angle, 10 degrees; echo time, 15.7 ms; repetition time, 423 ms; field of view,  $340 \times 340$  mm; slice thickness, 0.9 mm; matrix size,  $448 \times 448$  pixels. The patient was injected with gadolinium-diethylenetriamine pentaacetic acid (Gd-DTPA) in a dose of 0.1 mmol/kg at an amount of 15 mL, and the dynamic sequences were acquired at 6 timepoints, including 1 pre-enhanced and 5 post-enhanced (from 1 to 4.5 min after enhanced). In the imaging-subtype validation cohort, the T1-weighted images were obtained by GE scanners on a 1.5-Tesla magnet strength using a three-dimensional spoiled gradient-echo sequence with a gadolinium-based contrast agent. The spacing between slices of validation images ranged from 2 to 3 mm, the in-plane resolution ranged from 0.53 to 0.86 mm, the acquisition matrix was  $256 \times 192$ , and the flip angle was 10 degrees. The echo time and repetition time were not available in the validation cohort. The data used from the TCIA database in this study followed the Data Usage Policies and Restrictions of TCIA [26].

## 2.5. Tumor Segmentation

Firstly, to correct the motion during dynamic enhancement, the post-contrast images were registered to the pre-contrast images using the affine registration method. Then, the threshold segmentation method was used to obtain VOI masks of tumor lesions roughly from the subtracted images of the first enhanced sequences. Both image registration and segmentation were performed in the open-source software 3D Slicer. Then, two radiologists blinded to the clinical information, one with ten years and another with three years of breast imaging experience, corrected the masks manually and confirmed tumor masks consensually.

### 2.6. Voxel-Based PE and SER Map Calculation

In order to not just extract features from the MR images and explore more radiomics features by taking advantage of the DCE-MRI technique, we tried to define and calculate the voxel-based PE and SER maps of VOI based on the signal intensity of each voxel for each patient in this work, according to the Breast Imaging-Reporting and Data System (BI-RADS) and some previous studies [27–29]. The N4 bias correction algorithm was applied to avoid data heterogeneity bias in all the 3T-MR images [30]. Then, the voxel-based PE maps quantified the relative change in the signal intensity for each voxel before and after contrast enhancement; voxel-based SER maps can describe the comparison of the signal intensity for each voxel during the post-contrast period. In this study, DCE-MRI sequences at four timepoints, including pre-contrast, early, middle, and late post-contrast images, were used to calculate the PE or SER maps of VOI. The voxel-based early PE map for each sample was obtained by Equation (1):

$$PE_{early} = 100 \times \frac{I_{early} - I_{pre}}{I_{pre}} \quad (1)$$

where  $I_{early}$  is the signal intensity of each voxel in the early contrast, and  $I_{pre}$  is the voxel initial intensity before contrast. The voxel-based SER maps of both middle and late enhancement were calculated by Equation (2):

$$SER_{map} = \frac{I_{early} - I_{pre}}{I_{map} - I_{pre}} \quad (2)$$

where  $I_{map}$  here is the signal intensity of each enhanced voxel in middle or late images. As mentioned above, the effects of motion between different timepoints were eliminated because the post-contrast images were already registered on the pre-contrast images. N4 bias correction and the calculation of PE and SER maps were performed in Python 3.5.2.

### 2.7. Image Preprocessing and Radiomics Feature Extraction

Before feature extraction, images were resampled to a uniform voxel spacing (1 mm × 1 mm × 1 mm) by the B-Spline method and were normalized as well as remapped in the histogram to fit within  $\mu \pm 3\sigma$  ( $\mu$ : mean gray-level within the volume of segmentation;  $\sigma$ : gray-level standard deviation), as absolute signal intensity values were not necessarily comparable between scanners. Pyradiomics (version 2.2.0) was used to perform image normalization, resampling, and radiomics feature extraction [31]. According to the Image Biomarker Standardization Initiative (IBSI), 14 shape features, 18 first-order features, and 22 gray level co-occurrence matrix (GLCM) texture features were extracted from the pre-contrast MR images for each patient in this study. Similarly, 18 first-order features and 22 GLCM texture features were extracted from the early PE, middle SER, and late SER maps, respectively. In total, 174 radiomics features were calculated and used in this study. Image preprocessing and feature extraction were all performed in Python 3.5.2.

### 2.8. Identification and Analysis of Imaging Subtypes

The consensus clustering algorithm was used to identify intrinsic imaging subtypes in the discovery and validation cohorts [32]. The algorithm first subsamples both items and features from a data matrix and then clusters them into  $k$  classes. The process is repeated multiple times. The proportion of clustering runs in which two items are grouped in multiple repetitive clustering, named as pairwise consensus value, is calculated. The algorithm generates a consensus matrix for a given number of clusters  $k$ , which can provide a quantitative method to estimate the number of unsupervised classes in a dataset.

In this work, we scaled imaging features by z-score and performed a bootstrap procedure with 10,000 times 80% items and 80% features resampling using the partitioning around medoids algorithm with Spearman distance metric. By varying the number of clusters  $k$  from 2 to 8, we selected the optimal number of clusters, which generated the

most stable consensus matrices and the most unambiguous cluster assignments across clustering runs. The optimal number of intrinsic unsupervised clusters was determined in the discovery cohort, and the same procedure was performed on the imaging-subtype validation cohort. Additionally, in-group proportion (IGP) statistical analysis was also used to demonstrate the reproducibility and reliability of novel subtypes [33]. IGP will be 100% if the clusters are identical between two datasets and will be 0% conversely.

One-way analysis of variance (ANOVA) and Tukey's 'Honest Significant Difference' test was used to determine imaging features specific to each imaging subtype. The Student's *t*-test was used to compare the significant differences in imaging features between two imaging subtypes. The 95% confidence interval (CI) was used to assess the differences in DCE-MRI features of each subtype. Principal component analysis (PCA) was performed to identify the contribution of imaging features to the novel subtypes.

We applied two-sided Pearson's chi-squared test or Fisher's exact test to test the independence of imaging subtypes from other clinical characteristics, including ER, PR, HER2, and Ki67, IHC subtypes, PAM50 subtypes, and clinical stages. The consensus clustering, IGP, and statistical analysis were performed in R 4.0.1.

To explore the relationship between imaging subtypes based on radiomics features and the parameters of the conventional pharmacokinetic model, we further calculated two key DCE-MRI quantitative parameters,  $K^{\text{trans}}$  and  $v_e$ , in the discovery cohort using a population-averaged arterial input function (AIF)-based Tofts model [34,35].  $K^{\text{trans}}$  was calculated by measuring the accumulation of Gd-DTPA-based contrast agent in the extracellular-extravascular space, and  $v_e$  was the fractional volume for extracellular space [35]. Statistical differences in  $K^{\text{trans}}$  and  $v_e$  between different imaging subtypes were analyzed using ANOVA and Student's *t*-test. The estimation of pharmacokinetic parameters was implemented in 3D Slicer and statistical analysis was performed in R 4.0.1.

### 2.9. RNA Sequencing and Transcriptomic Analysis

Tumor tissue was collected from 199 patients in the discovery cohort, and the protocols of total RNA isolation and sequencing are described in the Supplementary Materials. Trimmomatic was used to control the sequencing quality with the following parameters: [LEADING:3 TRAILING:5 SLIDINGWINDOW: 4:15 MINLEN:60] [36]. RNA-seq reads were aligned to human genome 19 by STAR [37] and quantified by HTSeq-Count [38]. The expression values of 57,773 transcripts were quantified in the forms of both counts and FPKM (fragments per kilobase of exon per million reads mapped).

We identified PAM50 intrinsic subtypes for the patients in the discovery cohort by using *geneFu* [39] and performed differential expression analysis between two imaging subtypes using DESeq2 [40]. Genes with adjusted *p*-values less than 0.05 were considered differentially expressed genes (DEGs). ANOVA was used to obtain subtype-specific genes among three DCE-MRI subtypes. Gene set enrichment analysis (GSEA) was conducted based on the DEGs to identify the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that differ significantly between imaging subtypes with the cutoff of false discovery rate (FDR) less than 0.25 [41,42]. We used CIBERSORT with a BC-specific reference signature matrix developed from single-cell transcriptome data to characterize the relative proportion of the 15 cell types, including malignant, fibroblasts, and immune cells, in bulk RNA-seq data [43,44]. ANOVA was used to determine specific cell types for each imaging subtype. The Student's *t*-test was used to compare the significant differences in cell proportions between two imaging subtypes. All analysis was performed in R 4.0.1.

### 2.10. Prognostic Analysis

Kaplan–Meier analysis with the log-rank test was used to assess the differences in disease-free survival (DFS) or overall survival (OS) among imaging subtypes. Due to the lack of image data in the external prognosis validation datasets, we established a classifier using the random forest with default parameters, except the number of trees was set to 2000, to predict imaging subtypes from gene expression profiles according to the imaging

subtype-specific genes. The micro area under the receiver operating characteristic curve (AUC) for multi-classification was used to assess classifier performance. The analysis was performed in R 4.0.1.

### 3. Results

#### 3.1. Identification and Validation of the DCE-MRI Subtypes

The patient distribution of the two radiogenomics cohorts was not different except for the PAM50 molecular subtypes (Table 1), which indicated that there was no significant clinical bias for the patients. Then, we identified three de novo subtypes based on 174 DCE-MRI radiomics features using consensus clustering in the imaging-subtype discovery cohort (Figure 2a). The 3-cluster solution corresponded to the largest and optimal cluster number in the discovery cohort, which induced the least incremental change in the area under the cumulative distribution function (CDF) curve while keeping the maximal consensus within clusters and the minimal rate of ambiguity in cluster assignments (Figure 2b). Then, we independently applied the same analysis procedure on the imaging-subtype validation cohort, and, interestingly, we observed that samples in the validation cohort were also clustered into three optimal clusters based on the 174 features, similar to the discovery cohort (Figure 2c,d). The numbers of sample for imaging subtype 1, subtype 2 and subtype 3 were 50, 62, and 62 in the discovery cohort, respectively, as well as 15, 25, and 32 in the validation cohort. The reproducibility of the three imaging subtypes across both cohorts was evaluated by using IGP, resulting in values of 90.91%, 94%, and 90.91% for each subtype in the validation cohort. In short, the three DCE-MRI subtypes identified in the discovery cohort were validated in the validation cohort, indicating that these subtypes can reflect the intrinsic image heterogeneity of BC.

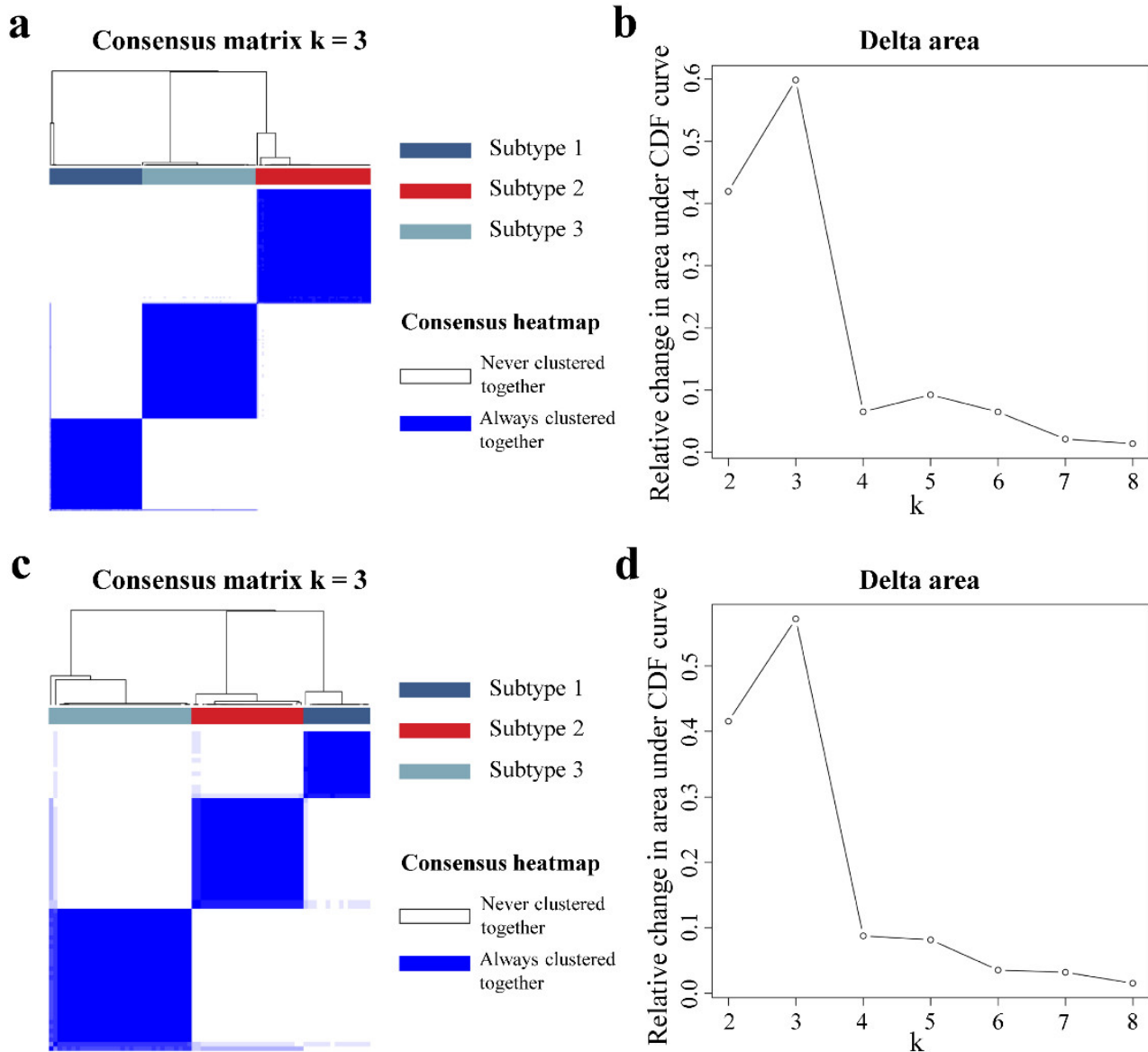
**Table 1.** Demographics of BC patients in the imaging-subtype discovery and validation cohorts. The detailed clinical, imaging, and molecular data for the discovery and validation of radiomics cohorts. The patient distribution of the two cohorts was not different except for the PAM50 molecular subtypes.

Characteristics	Discovery Cohort (n = 174)	Validation Cohort (n = 72)	p-Value
Age, mean (SD)	≤50 years:95/>50 years:79; 49.78 years (9.99)	≤50 years:30/>50 years:42; 53.96 years (11.75)	0.088 <sup>a</sup>
Histopathology type			
Ductal	156	71	0.069 <sup>a</sup>
Lobular	4	0	
Mixed	12	0	
Other	2	1	
BI-RADS			
Category 3	1	NA	NA
Category 4	61	NA	
Category 5	103	NA	
Category 6	9	NA	
IHC receptors			
ER status	P:127/N:47	P:61/N:11	0.071 <sup>a</sup>
PR status	P:111/N:63	P:55/N:17	0.077 <sup>a</sup>
HER2 status	P:36/N:138	P:14 / N:37/NA:21	0.407 <sup>a</sup>
Ki67 status	high:136/low:38	NA	NA
IHC-based subtype			
Luminal-A	28	NA	NA
Luminal-B	101	NA	NA
HER2-positive	15	NA	NA
Triple-negative	30	NA	NA

Table 1. Cont.

Characteristics	Discovery Cohort (n = 174)	Validation Cohort (n = 72)	p-Value
PAM50 subtype			
Luminal-A	49	44	<0.001 <sup>a</sup>
Luminal-B	43	9	
HER2-Enriched	29	5	
Basal-like	43	10	
Normal-like	10	4	
Pathological stage			
Stage I	55	17	0.267 <sup>a</sup>
Stage II	94	47	
Stage III	25	8	

Note: unless otherwise indicated, data are the number of patients or the p-value of the statistical test. <sup>a</sup> p-value for the two-sided Pearson's chi-squared test. SD, standard deviation; P, positive receptor status; N, negative. BI-RADS, Breast Imaging-Reporting and Data System. NA: not available.



**Figure 2.** Unsupervised consensus clustering of DCE-MRI features. In the imaging-subtype discovery (a) and validation cohorts (c), the consensus matrix heatmaps for the optimal cluster number (k = 3) are displayed. Rows and columns of the consensus matrix were samples and values ranging from zero



to one, indicating the two samples never clustered together or always clustered together, respectively. The dendrograms above the heatmaps indicate the samples ordering in three subtypes. The cluster number  $k$  varied from 2 to 8, and the optimal  $k$  could be determined to be 3 when yielding the largest relative change in area under the cumulative distribution function (CDF) curves for both the discovery (b) and validation cohorts (d).

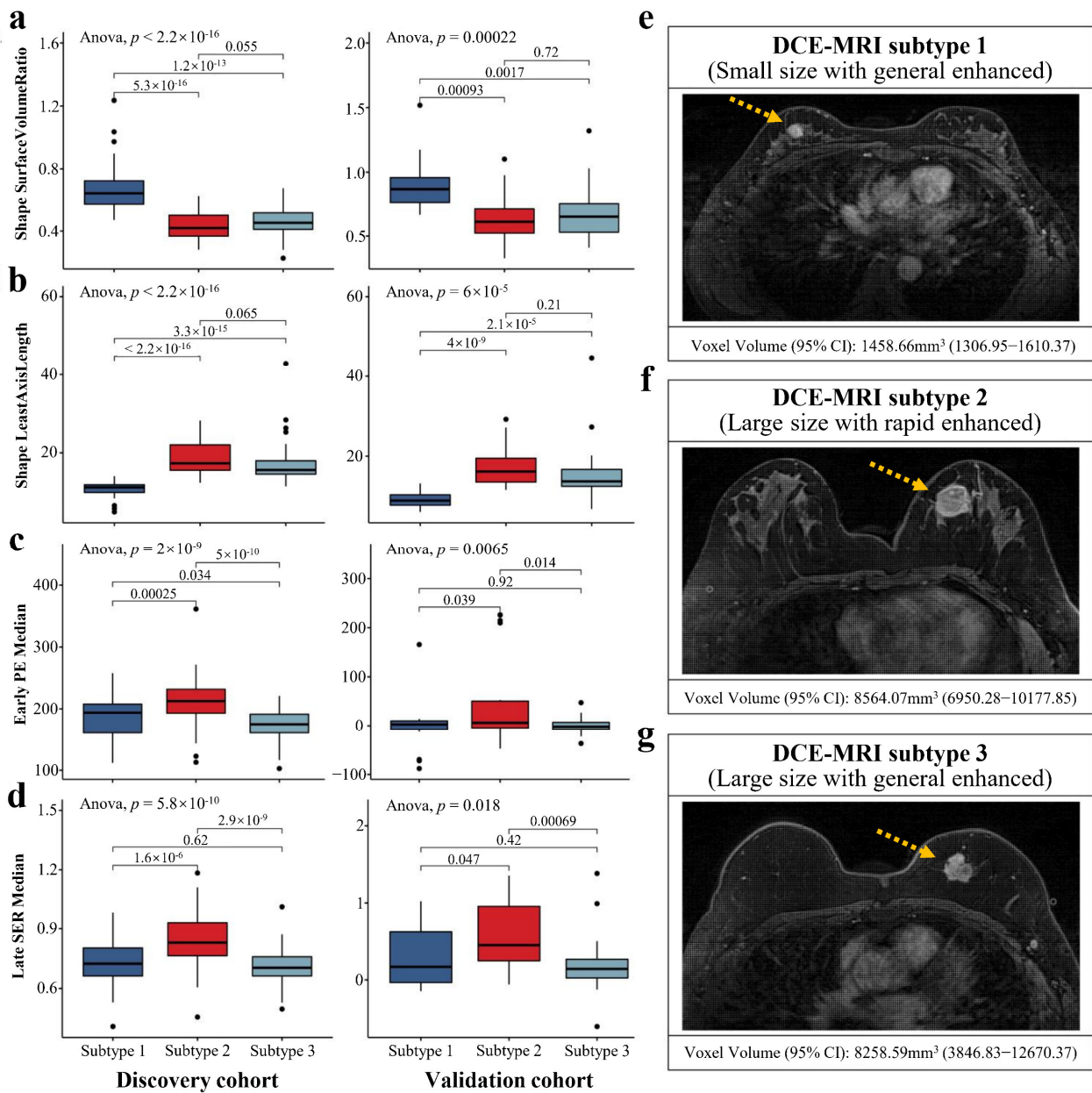
### 3.2. Imaging Characteristics of the DCE-MRI Subtypes

We examined the difference in imaging features among the DCE-MRI subtypes. The  $p$ -values of ANOVA for each imaging feature in the discovery and validation cohorts are listed in Table S1. The PCA results indicated that the 58 important imaging features constitute the heterogeneity of the three imaging subtypes (Figure S2). Remarkably, we found that tumor shape features, such as surface volume ratio and least axis length, showed extremely significant differences among the three imaging subtypes in both cohorts ( $p$ -value  $< 0.001$ , Figure 3a,b). In the discovery cohort, subtype 1 typically had smaller tumor size (mean voxel volume [95% CI], 1458.66 mm<sup>3</sup> [1306.95–1610.37]) compared to subtypes 2 and 3 ( $p$ -value  $< 0.01$ ), while there was no significant difference in tumor size between subtypes 2 (mean voxel volume [95% CI], 8564.07 mm<sup>3</sup> [6950.28–10,177.85]) and 3 (mean voxel volume [95% CI], 8258.59 mm<sup>3</sup> [3846.83–12,670.37]). The imaging features that can reflect dynamic enhanced characteristics of the tumor, including the medians of early PE and late SER maps, demonstrated highly significant differences among imaging subtypes in both cohorts ( $p$ -value  $< 0.05$ , Figure 3c,d). We found that during dynamic contrast enhancement, BC samples of subtype 2 showed a pattern of rapid enhancement, while subtypes 1 and 3 showed a generally enhanced pattern. Based on the clinical imaging characteristics, we respectively defined subtype 1 as small size with generally enhanced pattern tumor (Figure 3e), subtype 2 as large size with rapidly enhanced pattern tumor (Figure 3f), and subtype 3 as large size with generally enhanced pattern tumor (Figure 3g).

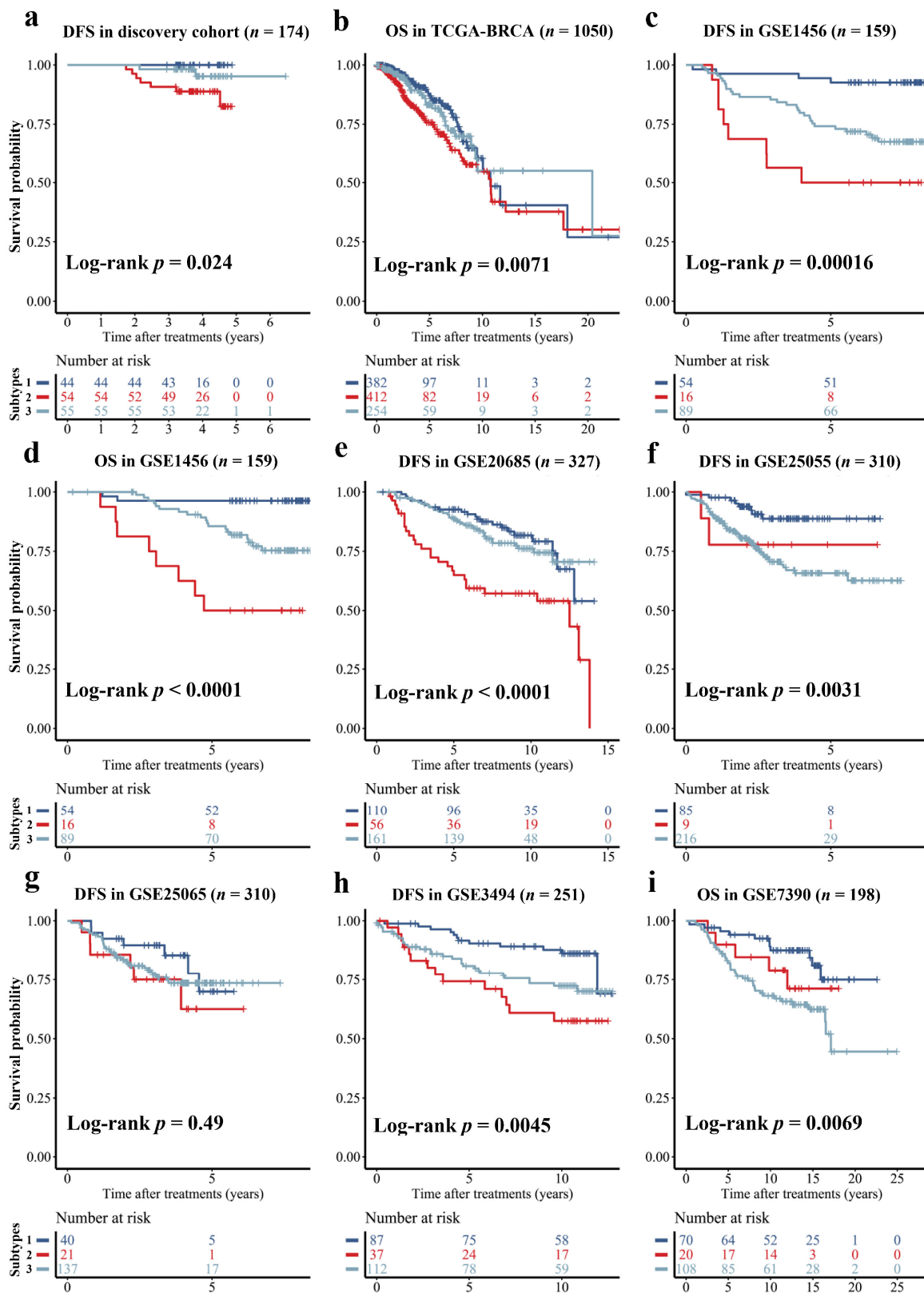
It was interesting to note that when analyzing the differences in pharmacokinetic parameters between the three imaging subtypes, we found significant differences in  $K^{\text{trans}}$  among the imaging subtypes ( $p$ -value  $< 0.01$ , Figure S3a). Compared to subtypes 1 and 3, BC patients with subtype 2 had significantly higher  $K^{\text{trans}}$  values (mean value, 0.1867 min<sup>-1</sup>), which was consistent with the findings based on the rapid enhancement pattern defined by quantitative radiomics. Although  $v_e$  did not differ significantly among the three imaging subtypes, it was different between subtypes 1 and 2 ( $p$ -value  $< 0.05$ , Figure S3b).

### 3.3. Distinct Prognostic Outcomes of the DCE-MRI Subtypes

We found significant DFS differences among the three DCE-MRI subtypes in the discovery cohort ( $p$ -value  $< 0.05$ , Figure 4a). To assess the prognostic ability of imaging subtypes further, we selected 993 subtype-specific genes by one-way ANOVA in the discovery cohort and established the classifier using random forest to predict imaging subtypes from gene expression profiles. This classifier achieved a moderate multi-classification AUC with 0.6 in the independent validation cohort for the prediction of imaging subtypes. Then, we applied the classifier to identify imaging subtypes in six external prognosis validation datasets which only had gene expression profiles, and found significantly different prognostic outcomes, including DFS and OS, for the imaging subtypes ( $p$ -value  $< 0.01$ , Figure 4). Notably, the pattern of outcome in external prognosis validation datasets was consistent with the imaging-subtype discovery cohort, with subtype 1 showing a favorable prognosis, subtype 2 having the worst prognosis, and subtype 3 displaying an intermediate prognosis.



**Figure 3.** Significant differences in tumor sizes and enhanced patterns shaped the imaging subtypes. Three subtypes were significantly associated with four representative quantitative imaging features in both cohorts, including surface volume ratio (a), least axis length (b), early PE map median (c), and late SER map median (d). ANOVA was used to identify the differences in imaging features among three subtypes, and Student’s *t*-test was used for the comparison between two subtypes. The mean values with 95% CI of the selected imaging feature and the representative sample from the discovery cohort for each imaging subtype are illustrated in (e–g) (yellow arrows point to the tumors).



**Figure 4.** Prognosis of DFS or OS stratified by the DCE-MRI subtypes. The Kaplan-Meier curve was plotted for the stratified samples by three imaging subtypes in the discovery cohort (a), and evaluated in seven external datasets without imaging data but by using a gene-profile-based imaging subtypes classifier to determine the imaging subtypes (b–i).

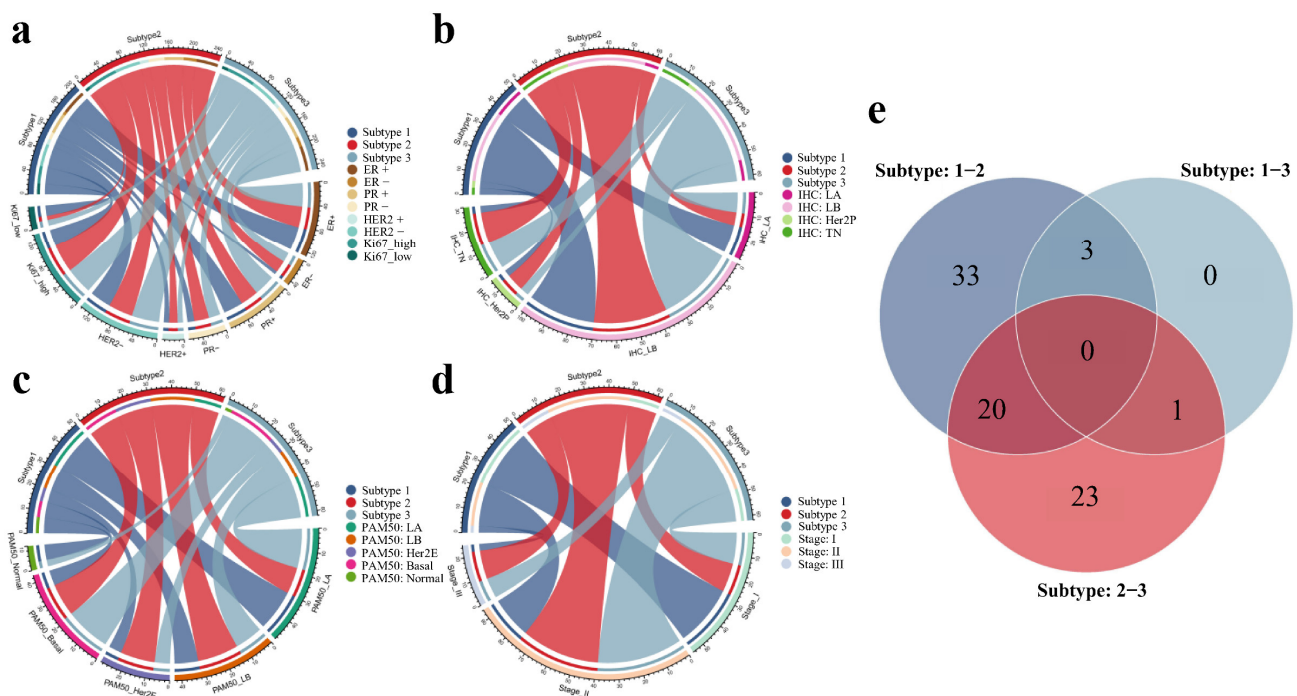
### 3.4. Associations with the Established BC Subtypes and Clinical Stages

Associations between the novel imaging subtypes and established BC subtypes or tumor stages in the discovery cohort are detailed in Table 2. We noticed imaging subtypes were significantly associated with Ki67 status, PAM50 intrinsic molecular subtypes, and tumor stages ( $p$ -value < 0.01). In the discovery cohort, 77.2% of samples with high Ki67 expression belonged to subtype 2 or subtype 3, whereas subtype 1 was more likely to have a low Ki67 expression ( $p$ -value = 0.004, Figure 5a). We also observed that 90% of TN samples belonged to subtype 2 or subtype 3 (Figure 5b), and a similar distribution (83.7%) was revealed in basal-like cases ( $p$ -value = 0.005, Figure 5c). Remarkably, we found that the proportion of favorable prognosis for subtype 1 became lower as the tumor stage increased, while subtypes 2 and 3 with worse prognosis were more prevalent in higher stage ( $p$ -value < 0.001, Figure 5d). High Ki67 expression, TN or basal-like subtype and high-grade tumor stage were usually regarded as the potential risk factors for clinical assessment of BC malignancy and poor prognosis. The distinct patterns of clinical outcome of imaging subtypes were consistent with the associations between the three subtypes and these risk factors.

**Table 2.** Differences of established subtypes and clinical stages for DCE-MRI subtypes in the imaging-subtype discovery cohort. The relationships between three imaging subtypes and established subtypes of BC, including clinical receptor-based, IHC-based and PAM50-based subtypes, as well as the clinical stages, were calculated.

Factors	Subtype 1 ( $n = 50$ )	Subtype 2 ( $n = 62$ )	Subtype 3 ( $n = 62$ )	$p$ -Value
ER status				
ER positive	44	39	44	0.011 <sup>a</sup>
ER negative	6	23	18	
PR status				
PR positive	37	34	40	0.109 <sup>a</sup>
PR negative	13	28	22	
HER2 status				
HER2 positive	10	16	10	0.409 <sup>a</sup>
HER2 negative	40	46	52	
Ki67 status				
Ki67 high	31	54	51	0.004 <sup>a</sup>
Ki67 low	19	8	11	
IHC-based subtype				
Luminal-A	13	6	9	0.051 <sup>b</sup>
Luminal-B	31	34	36	
HER2-positive	3	8	4	
Triple-negative	3	14	13	
PAM50 subtype				
Luminal-A	19	12	18	0.005 <sup>b</sup>
Luminal-B	11	19	13	
HER2-Enriched	6	16	7	
Basal-like	7	15	21	
Normal-like	7	0	3	
Pathological stage				
Stage I	28	16	14	<0.001 <sup>a</sup>
Stage II	19	36	39	
Stage III	3	13	9	

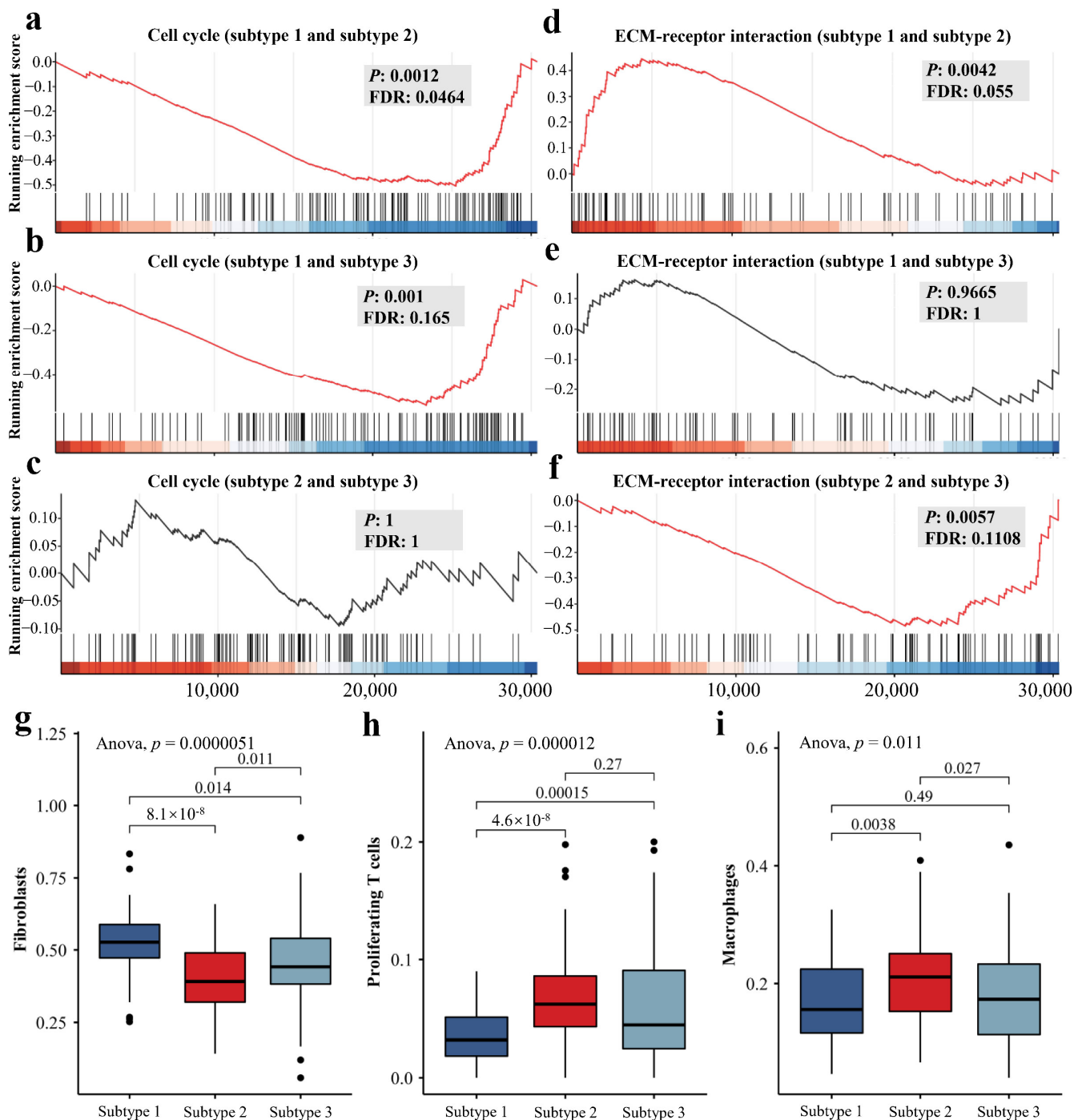
Note: unless otherwise indicated, data are number of patients or the  $p$ -value of the statistical test. <sup>a</sup>  $p$ -value for the two-sided Pearson's chi-squared test, <sup>b</sup>  $p$ -value for the two-sided Fisher's exact test.



**Figure 5.** Associations of the DCE-MRI subtypes with clinical and pathway characteristics. The chord diagrams between three imaging subtypes and clinical IHC receptors (a), IHC-based subtypes (b), PAM50 molecular subtypes (c), and clinical stages (d) in the discovery cohort were described. Venn diagram of the differentially enriched KEGG pathways between two of three imaging subtypes (e).

### 3.5. The Differences of Molecular and Microenvironment Characteristics among Imaging Subtypes

We identified 56 different enriched KEGG pathways between subtypes 1 and 2, 4 pathways between subtypes 1 and 3, and 44 between subtypes 2 and 3, respectively, by gene differential expression analysis and GSEA (FDR < 0.25, Tables S2–S4). These imaging subtype-associated KEGG pathways included cell cycle, ECM-receptor interaction, Hedgehog signaling pathway, proteoglycans in cancer, PI3K-Akt signaling pathway, Ras signaling pathway, and breast cancer. Interestingly, based on the heterogeneity of molecular pathways, subtype 2 was distant from the other two subtypes, while subtypes 1 and 3 were much closer (Figure 5e). Compared to subtype 1, the cell cycle was significantly enriched at the bottom in both subtypes 2 and 3 (FDR = 0.0464 and 0.165, Figure 6a,b), while there was no difference between subtype 2 and subtype 3 (Figure 6c). We also observed that ECM-receptor interaction was significantly different between imaging subtype 2 and subtypes 1 and 3 (FDR = 0.055 and 0.1108, Figure 6d,f), while there was no difference between subtype 1 and subtype 3 (Figure 6e). Notably, the enrichment trend of ECM-receptor interaction in subtype 2 with a worse prognosis was similar, compared to the subtypes with slightly better prognosis. The extracellular matrix (ECM) is an essential component of the tumor microenvironment and plays an important role in tumorigenesis, proliferation, and metastasis. Therefore, we further estimated the abundance of 15 cell types in the samples from both discovery and validation cohorts, including malignant, fibroblasts, and immune cells. Significant differences in the abundance of fibroblasts, proliferating T cells and macrophages were revealed among three imaging subtypes in the discovery cohort ( $p$ -value < 0.05, Figure 6g–i, Table 3), and interestingly similar trends were also observed in the independent validation cohort (Figure S4).



**Figure 6.** Important enriched KEGG pathways and tumor microenvironment characteristics of different DCE-MRI subtypes. Detailed differences in representative transcriptional behavior among three subtypes including cell cycle (a–c) and ECM-receptor interaction (d–f) are illustrated. The boxplots for the fractions of cancer-associated fibroblasts (g), proliferating T cells (h) and macrophages (i) among imaging subtypes in both discovery and validation cohorts are displayed. ANOVA was used to identify the differences in cellular fractions among three subtypes, and Student’s t-test was used for the comparison between two subtypes.

**Table 3.** Immune microenvironment cellular fraction differences of imaging subtypes in the imaging-subtype discovery cohort. A BC-specific reference signature matrix developed from single-cell transcriptome data was used to estimate the relative proportion of 15 cell types in the bulk RNA-seq data. Then, the differences in the cellular fraction of imaging subtypes were assessed.

Cell Types	<i>p</i> -Value	Subtype 2-1	Subtype 3-1	Subtype 3-2
Malignant cells	0.166	0.141	0.517	0.67
Fibroblasts	$5.11 \times 10^{-6}$ *	$2.58 \times 10^{-6}$ *	0.031*	0.021 *
Proliferating T cells	$1.18 \times 10^{-5}$ *	$1.06 \times 10^{-5}$ *	0.001*	0.416
Cytotoxic T cells	NaN	NaN	NaN	NaN
Regulatory T cells	0.738	0.751	0.795	0.996
Naive-like T cells	0.408	1	0.504	0.464
Natural killer cells	0.408	1	0.504	0.464
Neutrophils	0.697	0.848	0.969	0.683
Plasma cells	0.241	0.218	0.759	0.565
Dendritic cells	0.408	0.504	1	0.464
Macrophages	0.011 *	0.013 *	0.773	0.059
Monocytes	0.8	0.911	0.976	0.788
Mast cells	0.408	0.504	1	0.464
B cells	0.704	0.896	0.679	0.91
Transitional T cells	0.884	0.892	0.996	0.923

Note: unless otherwise indicated, data are the *p*-value of ANOVA or Tukey's test. \* *p*-value for *p*-value < 0.05. NaN: Not a Number.

#### 4. Discussion

In this study, we identified and validated three distinct BC imaging subtypes by using DCE-MRI features of tumor lesions in two cohorts. The novel BC subtypes had significantly different tumor sizes and imaging enhancement patterns, as well as clinical outcomes. Importantly, the independent prognostic value of imaging subtypes was demonstrated in our discovery cohort, and by using a gene expression-based subtype classifier, we displayed this value in six external datasets. We further discovered that the imaging subtypes significantly correlated with Ki67, PAM50 subtypes and clinical stages, and revealed the biological mechanisms underlying this heterogeneity. Three subtypes not only differed significantly in their transcriptional activities such as cell cycle and ECM-receptor interaction KEGG pathways, but also had different microenvironment characteristics, particularly in the abundance of fibroblasts, proliferating T-cells, and macrophages. In summary, we revealed BC heterogeneity from a new perspective and uncovered the possible biological mechanisms, which may be useful for BC clinical decision making.

Tumor size and enhancement patterns were the most representative imaging differences in our imaging subtypes. As a critical determinant of clinical outcome and staging system for BC, tumor size is strongly correlated with prognosis, and a larger BC usually suggests a worse outcome. A tumor with a rapid uptake in the early-enhanced process of DCE-MRI and a quick washout in the later part of enhancement suggests more malignancy [45–47]. Imaging subtype 2 with larger tumor size, early enhanced rapid pattern and quickly reduced late pattern (Figures 3b–d and S3a) were significantly associated with poorer DFS of BC (Figure 4a). Conversely, subtype 1 had the best clinical outcome, indicating that our findings were consistent with previous studies. More importantly, the three subtypes displayed high reproducibility (IGP over 90%) in the two cohorts (Figure 2) and had similar prognostic trends in multiple external datasets (Figure 4).

The heterogeneity of imaging features may arise from the differences in biological mechanisms. We identified 80 out of 186 KEGG pathways associated with the imaging subtypes, suggesting a link between DCE-MRI features and transcriptional activities. These pathways, including cell cycle, ECM-receptor interaction, PI3K-Akt signaling pathway and homologous recombination, fall into several broad categories involving a variety of BC-related biological activities such as cellular processes, metabolism, and genetic and environmental information processing. We found biologically significant differences between imaging subtype 2 and other subtypes, while subtypes 1 and 3 were more similar. The

tail-end enrichment of cell cycle and a higher proportion of Ki67 low-expressing patients, implied a reduced capacity for tumor growth and proliferation in subtype 1, resulting in smaller tumor size and favorable prognosis. Profound changes in cellular metabolism were usually a hallmark of malignancy, and subtype 2, which had the worst prognosis displayed remarkable metabolic disorders compared to other subtypes [48–50]. For example, as a key oxidative enzyme, cytochrome P450 can metabolize many carcinogens and anticancer drugs in BC [51]. We discovered that two cytochrome P450-related pathways (metabolism of xenobiotics by cytochrome P450 and drug metabolism-cytochrome P450) were specific to subtype 2. Furthermore, although both subtypes 2 and 3 had larger tumor size (Figure 3a,b), their enhancement patterns were significantly different (Figures 3c,d and S3a). The possible biological mechanisms for this result could be explained by the fact that subtypes 2 and 3 displayed no difference in cell cycle but had significant differences in many pathways, such as ECM-receptor interaction and some metabolic pathways. ECM is an essential component of the tumor microenvironment and plays an important role in regulating BC progression and metastasis, which might affect the process of enhancement imaging [52–55]. Our work may provide a non-invasive perspective on the biology of BC.

Differences in the tumor immune microenvironment emerged in the imaging subtypes. Subtype 2 exhibited a lower proportion of fibroblasts and a higher proportion of proliferating T cells and macrophages. Cancer-associated fibroblasts (CAFs) are not only important components of the BC microenvironment and are essential in tumor growth and development, but also have complex phenotypes and functional heterogeneity [56–59]. T cell proliferation is closely related to immune response, treatment resistance, and prognosis of BC [60–62]. Macrophages are important players in the extracellular matrix activity of breast tumors and a high fraction of macrophages usually suggests a poorer prognosis [63,64].

Differing from most radiogenomics studies that focused on the association of imaging features with gene expression profiles, we uncovered the heterogeneity of BC from the novel perspective of radiomics, and further revealed the differences of clinical outcome and biological activities for imaging subtypes. Although some previous works have used similar approaches to identify potential imaging subtypes of glioblastoma and BC based on MRI data [65,66], our work still has some strengths, especially compared to the work of Wu et al. [66]. Firstly, we identified three imaging subtypes of BC in a larger dataset and achieved high cluster reproducibility in an independent multi-center validation cohort (all of our IGPs were greater than 90%, whereas only one IGP of Wu et al. [66] was greater than 90%). Secondly, the determination of breast background parenchymal tissue is a challenging task in clinical practice, whereas the segmentation of tumor lesions is much easier. We predicted the prognosis well using only the information of tumor lesions, which is more convenient for clinical applications. Additionally, we highlighted the clinical evidence of DCE-MRI features as non-invasive biomarkers. In the radiomics discovery cohort, the mean voxel volume with 95% CI of imaging subtype 1 was  $1458.66 \text{ mm}^3$  [1306.95–1610.37], and larger tumor sizes were observed in subtype 2 ( $8564.07 \text{ mm}^3$  [6950.28–10,177.85]) and subtype 3 ( $8258.59 \text{ mm}^3$  [3846.83–12,670.37]). Finally, we not only revealed the differences in biological functions of imaging subtypes, but also further analyzed the associations of cellular fractions with imaging subtypes, although the abundance of cells was estimated from bulk RNA-seq data. It was noteworthy that we used a BC-specific expression matrix based on single-cell data to estimate cellular fractions in bulk RNA-seq data, which indicated that the tumor microenvironment was disease-specific. There were some limitations in our work. Studies supporting the possible influence of biological activity on imaging features, and whether this correlation is causal, still need to be verified in conjunction with biological experiments. The enhancement patterns of DCE-MR imaging that define imaging subtypes need to be analyzed for differences in vascular permeability based on histopathology data to further validate our findings. The relationship between the complex characteristics of the tumor microenvironment and imaging subtypes may require deeper mining in combination with single-cell RNA-seq data. Additionally, the gene-profile-based classifier of DCE-MRI subtypes did not perform very well, and should be improved based on more



radiogenomics cohorts of BC in the future, and the robustness of the prognostic value of the imaging features needs to be further evaluated in larger datasets for clinical applicability.

## 5. Conclusions

In this study, three novel BC subtypes were identified and validated based on the radiomics features from tumor lesions in two independent DCE-MRI cohorts. The imaging subtypes showed distinct clinical outcomes, molecular characteristics, and cellular fractions in the tumor microenvironment. Our work may provide a potential radiomics biomarker for the precision medicine of BC.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers14225507/s1>, Figure S1: The inclusion and exclusion criteria of BC patients for radiomics cohorts. Figure S2: PCA of significantly different imaging features in the imaging-subtype discovery cohort. Figure S3: Differences in pharmacokinetic parameters of patients with different imaging subtypes in the imaging-subtype discovery cohort. Figure S4: Distinct cellular fractions in the tumor microenvironment of different imaging subtypes in the imaging-subtype validation cohort. Table S1: The *p*-values of ANOVA for DCE-MR features of three imaging subtypes in the imaging-subtype discovery and validation cohorts. Table S2: Differentially enriched KEGG pathways between imaging subtypes 1 and 2 in the imaging-subtype discovery cohort. Table S3: Differentially enriched KEGG pathways between imaging subtypes 1 and 3 in the imaging-subtype discovery cohort. Table S4: Differentially enriched KEGG pathways between imaging subtypes 2 and 3 in the imaging-subtype discovery cohort.

**Author Contributions:** Conceptualization, W.M., X.S. and H.L.; Data curation, F.L., Y.Z., Y.L. and X.L.; Formal analysis, W.M.; Funding acquisition, Y.L., X.S. and H.L.; Investigation, F.L., Y.Z. and Y.B.; Methodology, W.M., Y.B. and W.G.; Project administration, X.L., X.S. and H.L.; Resources, Y.Z., Y.B., W.G., Y.L., X.L. and H.L.; Software, F.L. and W.G.; Supervision, X.S. and H.L.; Validation, F.L.; Visualization, W.M. and F.L.; Writing—original draft, W.M.; Writing—review and editing, W.M., X.S. and H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the National Natural Science Foundation of China (81830053, 61972084, 61871121), the National Key R&D Program of China (2018YFC1314900, 2018YFC1314902), and the Bethune Charitable Foundation (G-X-2019-0101-12).

**Institutional Review Board Statement:** This study was approved by the Ethics Committees of Nanjing Medical University (2019SR512).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The RNA-seq data of the imaging-subtype discovery cohort are available at <https://ngdc.cnbc.ac.cn/bioproject/> (BioProject number: PRJCA005965, GSA-Human number: HRA001100, accessed on 1 October 2019). The TCGA-BRCA gene expression data are available at <https://portal.gdc.cancer.gov/projects/TCGA-BRCA> (accessed on 31 January 2019) and the corresponding DCE-MRI data are available at <https://wiki.cancerimagingarchive.net/display/Public/TCGA-BRCA> (accessed on 20 October 2019). Further information and other data that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** We thank Wenjing Cui from the Department of Radiology, Jiangsu Province Hospital of Chinese Medicine for her contribution to tumor segmentation.

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

## References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. Polyak, K. Heterogeneity in breast cancer. *J. Clin. Investig.* **2011**, *121*, 3786–3788. [[CrossRef](#)] [[PubMed](#)]
3. Martelotto, L.G.; Ng, C.K.Y.; Piscuoglio, S.; Weigelt, B.; Reis-Filho, J.S. Breast cancer intra-tumor heterogeneity. *Breast Cancer Res.* **2014**, *16*, 210. [[CrossRef](#)] [[PubMed](#)]

4. Perou, C.M.; Sørlie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; Rees, C.A.; Pollack, J.R.; Ross, D.T.; Johnsen, H.; Akslen, L.A.; et al. Molecular portraits of human breast tumours. *Nature* **2000**, *406*, 747–752. [[CrossRef](#)]
5. Sørlie, T.; Perou, C.M.; Tibshirani, R.; Aas, T.; Geisler, S.; Johnsen, H.; Hastie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10869. [[CrossRef](#)]
6. Hammond, M.E.H.; Hayes, D.F.; Dowsett, M.; Allred, D.C.; Hagerty, K.L.; Badve, S.; Fitzgibbons, P.L.; Francis, G.; Goldstein, N.S.; Hayes, M.; et al. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Immunohistochemical Testing of Estrogen and Progesterone Receptors in Breast Cancer. *J. Clin. Oncol.* **2010**, *28*, 2784–2795. [[CrossRef](#)]
7. Dowsett, M.; Nielsen, T.O.; A'Hern, R.; Bartlett, J.; Coombes, R.C.; Cuzick, J.; Ellis, M.; Henry, N.L.; Hugh, J.C.; Lively, T.; et al. Assessment of Ki67 in Breast Cancer: Recommendations from the International Ki67 in Breast Cancer Working Group. *JNCI J. Natl. Cancer Inst.* **2011**, *103*, 1656–1664. [[CrossRef](#)]
8. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; de Jong, E.E.C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [[CrossRef](#)]
9. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2015**, *278*, 563–577. [[CrossRef](#)]
10. Huynh, E.; Hosny, A.; Guthrie, C.; Bitterman, D.S.; Petit, S.F.; Haas-Kogan, D.A.; Kann, B.; Aerts, H.J.W.L.; Mak, R.H. Artificial intelligence in radiation oncology. *Nat. Rev. Clin. Oncol.* **2020**, *17*, 771–781. [[CrossRef](#)]
11. Aerts, H.J.W.L.; Velazquez, E.R.; Leijenaar, R.T.H.; Parmar, C.; Grossmann, P.; Carvalho, S.; Bussink, J.; Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **2014**, *5*, 4006. [[CrossRef](#)]
12. Pinker, K.; Chin, J.; Melsaether, A.N.; Morris, E.A.; Moy, L. Precision Medicine and Radiogenomics in Breast Cancer: New Approaches toward Diagnosis and Treatment. *Radiology* **2018**, *287*, 732–747. [[CrossRef](#)]
13. Xu, Z.; Wang, Y.; Chen, M.; Zhang, Q. Multi-region radiomics for artificially intelligent diagnosis of breast cancer using multimodal ultrasound. *Comput. Biol. Med.* **2022**, *149*, 105920. [[CrossRef](#)]
14. Wang, X.; Wang, S.; Yin, X.; Zheng, Y. MRI-based radiomics distinguish different pathological types of hepatocellular carcinoma. *Comput. Biol. Med.* **2022**, *141*, 105058. [[CrossRef](#)]
15. Nazari, M.; Shiri, I.; Zaidi, H. Radiomics-based machine learning model to predict risk of death within 5-years in clear cell renal cell carcinoma patients. *Comput. Biol. Med.* **2021**, *129*, 104135. [[CrossRef](#)]
16. Gong, L.; Xu, M.; Fang, M.; He, B.; Li, H.; Fang, X.; Dong, D.; Tian, J. The potential of prostate gland radiomic features in identifying the Gleason score. *Comput. Biol. Med.* **2022**, *144*, 105318. [[CrossRef](#)]
17. Chen, X.; Zhou, M.; Wang, Z.; Lu, S.; Chang, S.; Zhou, Z. Immunotherapy treatment outcome prediction in metastatic melanoma through an automated multi-objective delta-radiomics model. *Comput. Biol. Med.* **2021**, *138*, 104916. [[CrossRef](#)]
18. Huang, Y.-Q.; Liang, C.-H.; He, L.; Tian, J.; Liang, C.-S.; Chen, X.; Ma, Z.-L.; Liu, Z.-Y. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J. Clin. Oncol.* **2016**, *34*, 2157–2164. [[CrossRef](#)]
19. Fan, M.; Xia, P.; Clarke, R.; Wang, Y.; Li, L. Radiogenomic signatures reveal multiscale intratumour heterogeneity associated with biological functions and survival in breast cancer. *Nat. Commun.* **2020**, *11*, 4861. [[CrossRef](#)]
20. Yamamoto, S.; Maki, D.D.; Korn, R.L.; Kuo, M.D. Radiogenomic Analysis of Breast Cancer Using MRI: A Preliminary Study to Define the Landscape. *Am. J. Roentgenol.* **2012**, *199*, 654–663. [[CrossRef](#)]
21. Yamamoto, S.; Han, W.; Kim, Y.; Du, L.; Jamshidi, N.; Huang, D.; Kim, J.H.; Kuo, M.D. Breast Cancer: Radiogenomic Biomarker Reveals Associations among Dynamic Contrast-enhanced MR Imaging, Long Noncoding RNA, and Metastasis. *Radiology* **2015**, *275*, 384–392. [[CrossRef](#)] [[PubMed](#)]
22. Li, H.; Zhu, Y.; Burnside, E.S.; Huang, E.; Drukker, K.; Hoadley, K.A.; Fan, C.; Conzen, S.D.; Zuley, M.; Net, J.M.; et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *Npj Breast Cancer* **2016**, *2*, 16012. [[CrossRef](#)]
23. Bismeyer, T.; van der Velden, B.H.M.; Canisius, S.; Lips, E.H.; Loo, C.E.; Viergever, M.A.; Wesseling, J.; Gilhuijs, K.G.A.; Wessels, L.F.A. Radiogenomic Analysis of Breast Cancer by Linking MRI Phenotypes with Tumor Gene Expression. *Radiology* **2020**, *296*, 277–287. [[CrossRef](#)] [[PubMed](#)]
24. Grimm, L.J.; Mazurowski, M.A. Breast Cancer Radiogenomics: Current Status and Future Directions. *Acad. Radiol.* **2020**, *27*, 39–46. [[CrossRef](#)] [[PubMed](#)]
25. Ming, W.; Zhu, Y.; Bai, Y.; Gu, W.; Li, F.; Hu, Z.; Xia, T.; Dai, Z.; Yu, X.; Li, H.; et al. Radiogenomics analysis reveals the associations of dynamic contrast-enhanced-MRI features with gene expression characteristics, PAM50 subtypes, and prognosis of breast cancer. *Front. Oncol.* **2022**, *12*, 943326. [[CrossRef](#)]
26. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J. Digit. Imaging* **2013**, *26*, 1045–1057. [[CrossRef](#)]

27. Mazurowski, M.A.; Zhang, J.; Grimm, L.J.; Yoon, S.C.; Silber, J.I. Radiogenomic Analysis of Breast Cancer: Luminal B Molecular Subtype Is Associated with Enhancement Dynamics at MR Imaging. *Radiology* **2014**, *273*, 365–372. [[CrossRef](#)]
28. Burnside, E.S.; Sickles, E.A.; Bassett, L.W.; Rubin, D.L.; Lee, C.H.; Ikeda, D.M.; Mendelson, E.B.; Wilcox, P.A.; Butler, P.F.; D’Orsi, C.J. The ACR BI-RADS<sup>®</sup> Experience: Learning From History. *J. Am. Coll. Radiol.* **2009**, *6*, 851–860. [[CrossRef](#)]
29. Xiao, J.; Rahbar, H.; Hippe, D.S.; Rendi, M.H.; Parker, E.U.; Shekar, N.; Hirano, M.; Cheung, K.J.; Partridge, S.C. Dynamic contrast-enhanced breast MRI features correlate with invasive breast cancer angiogenesis. *Npj Breast Cancer* **2021**, *7*, 42. [[CrossRef](#)]
30. 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](#)]
31. 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.J.W.L. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* **2017**, *77*, e104. [[CrossRef](#)]
32. Monti, S.; Tamayo, P.; Mesirov, J.; Golub, T. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Mach. Learn.* **2003**, *52*, 91–118. [[CrossRef](#)]
33. Kapp, A.V.; Tibshirani, R. Are clusters found in one dataset present in another dataset? *Biostatistics* **2007**, *8*, 9–31. [[CrossRef](#)]
34. Tofts, P.S. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J. Magn. Reson. Imaging* **1997**, *7*, 91–101. [[CrossRef](#)]
35. Sourbron, S.P.; Buckley, D.L. On the scope and interpretation of the Tofts models for DCE-MRI. *Magn. Reson. Med.* **2011**, *66*, 735–745. [[CrossRef](#)]
36. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
37. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [[CrossRef](#)]
38. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)]
39. Gendoo, D.M.A.; Ratanasirigulchai, N.; Schröder, M.S.; Paré, L.; Parker, J.S.; Prat, A.; Haibe-Kains, B. Genefu: An R/Bioconductor package for computation of gene expression-based signatures in breast cancer. *Bioinformatics* **2016**, *32*, 1097–1099. [[CrossRef](#)]
40. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)]
41. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545. [[CrossRef](#)]
42. Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y. clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* **2012**, *16*, 284–287. [[CrossRef](#)]
43. Newman, A.M.; Liu, C.L.; Green, M.R.; Gentles, A.J.; Feng, W.; Xu, Y.; Hoang, C.D.; Diehn, M.; Alizadeh, A.A. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **2015**, *12*, 453–457. [[CrossRef](#)]
44. Li, H.; Huang, Y.; Sharma, A.; Ming, W.; Luo, K.; Gu, Z.; Sun, X.; Liu, H. From Cellular Infiltration Assessment to a Functional Gene Set-Based Prognostic Model for Breast Cancer. *Front. Immunol.* **2021**, *12*, 751530. [[CrossRef](#)] [[PubMed](#)]
45. Macura, K.J.; Ouwerkerk, R.; Jacobs, M.A.; Bluemke, D.A. Patterns of Enhancement on Breast MR Images: Interpretation and Imaging Pitfalls. *RadioGraphics* **2006**, *26*, 1719–1734. [[CrossRef](#)] [[PubMed](#)]
46. Moon, M.; Cornfeld, D.; Weinreb, J. Dynamic Contrast-Enhanced Breast MR Imaging. *Magn. Reson. Imaging Clin. North Am.* **2009**, *17*, 351–362. [[CrossRef](#)]
47. Tuncbilek, N.; Tokatli, F.; Altaner, S.; Sezer, A.; Türe, M.; Omurlu, I.K.; Temizoz, O. Prognostic value DCE-MRI parameters in predicting factor disease free survival and overall survival for breast cancer patients. *Eur. J. Radiol.* **2012**, *81*, 863–867. [[CrossRef](#)]
48. Boroughs, L.K.; DeBerardinis, R.J. Metabolic pathways promoting cancer cell survival and growth. *Nat. Cell Biol.* **2015**, *17*, 351–359. [[CrossRef](#)] [[PubMed](#)]
49. Pavlova Natalya, N.; Thompson Craig, B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab.* **2016**, *23*, 27–47. [[CrossRef](#)] [[PubMed](#)]
50. Lane, A.N.; Tan, J.; Wang, Y.; Yan, J.; Higashi, R.M.; Fan, T.W.M. Probing the metabolic phenotype of breast cancer cells by multiple tracer stable isotope resolved metabolomics. *Metab. Eng.* **2017**, *43*, 125–136. [[CrossRef](#)]
51. Murray, G.I.; Patimalla, S.; Stewart, K.N.; Miller, I.D.; Heys, S.D. Profiling the expression of cytochrome P450 in breast cancer. *Histopathology* **2010**, *57*, 202–211. [[CrossRef](#)]
52. Bao, Y.; Wang, L.; Shi, L.; Yun, F.; Liu, X.; Chen, Y.; Chen, C.; Ren, Y.; Jia, Y. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell Mol. Biol. Lett.* **2019**, *24*, 38. [[CrossRef](#)]
53. Jallow, F.; O’Leary, K.A.; Rugowski, D.E.; Guerrero, J.F.; Ponik, S.M.; Schuler, L.A. Dynamic interactions between the extracellular matrix and estrogen activity in progression of ER+ breast cancer. *Oncogene* **2019**, *38*, 6913–6925. [[CrossRef](#)]
54. Bahcecioglu, G.; Yue, X.; Howe, E.; Guldner, I.; Stack, M.S.; Nakshatri, H.; Zhang, S.; Zorlutuna, P. Aged Breast Extracellular Matrix Drives Mammary Epithelial Cells to an Invasive and Cancer-Like Phenotype. *Adv. Sci.* **2021**, *8*, 2100128. [[CrossRef](#)]
55. Cox, T.R. The matrix in cancer. *Nat. Rev. Cancer* **2021**, *21*, 217–238. [[CrossRef](#)]
56. Yu, Y.; Xiao, C.H.; Tan, L.D.; Wang, Q.S.; Li, X.Q.; Feng, Y.M. Cancer-associated fibroblasts induce epithelial–mesenchymal transition of breast cancer cells through paracrine TGF- $\beta$  signalling. *Br. J. Cancer* **2014**, *110*, 724–732. [[CrossRef](#)]

57. Costa, A.; Kieffer, Y.; Scholer-Dahirel, A.; Pelon, F.; Bourachot, B.; Cardon, M.; Sirven, P.; Magagna, I.; Fuhrmann, L.; Bernard, C.; et al. Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **2018**, *33*, 463–479. [\[CrossRef\]](#)
58. Becker, L.M.; O'Connell, J.T.; Vo, A.P.; Cain, M.P.; Tampe, D.; Bizarro, L.; Sugimoto, H.; McGow, A.K.; Asara, J.M.; Lovisa, S.; et al. Epigenetic Reprogramming of Cancer-Associated Fibroblasts Deregulates Glucose Metabolism and Facilitates Progression of Breast Cancer. *Cell Rep.* **2020**, *31*, 107701. [\[CrossRef\]](#)
59. Yu, S.; Lu, Y.; Su, A.; Chen, J.; Li, J.; Zhou, B.; Liu, X.; Xia, Q.; Li, Y.; Li, J.; et al. A CD10-OGP Membrane Peptolytic Signaling Axis in Fibroblasts Regulates Lipid Metabolism of Cancer Stem Cells via SCD1. *Adv. Sci.* **2021**, *8*, 2101848. [\[CrossRef\]](#)
60. Nagalla, S.; Chou, J.W.; Willingham, M.C.; Ruiz, J.; Vaughn, J.P.; Dubey, P.; Lash, T.L.; Hamilton-Dutoit, S.J.; Bergh, J.; Sotiropoulos, C.; et al. Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. *Genome Biol.* **2013**, *14*, R34. [\[CrossRef\]](#)
61. Ali, H.R.; Provenzano, E.; Dawson, S.J.; Blows, F.M.; Liu, B.; Shah, M.; Earl, H.M.; Poole, C.J.; Hiller, L.; Dunn, J.A.; et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12 439 patients. *Ann. Oncol.* **2014**, *25*, 1536–1543. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Savas, P.; Virassamy, B.; Ye, C.; Salim, A.; Mintoff, C.P.; Caramia, F.; Salgado, R.; Byrne, D.J.; Teo, Z.L.; Dushyanthen, S.; et al. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat. Med.* **2018**, *24*, 986–993. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Obeid, E.; Nanda, R.; Fu, Y.-X.; Olopade, O.I. The role of tumor-associated macrophages in breast cancer progression (review). *Int. J. Oncol.* **2013**, *43*, 5–12. [\[CrossRef\]](#)
64. Choi, J.; Gyamfi, J.; Jang, H.; Koo, J.S. The role of tumor-associated macrophage in breast cancer biology. *Histol. Histopathol.* **2018**, *33*, 133–145. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Itakura, H.; Achrol, A.S.; Mitchell, L.A.; Loya, J.J.; Liu, T.; Westbroek, E.M.; Feroze, A.H.; Rodriguez, S.; Echegaray, S.; Azad, T.D.; et al. Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities. *Sci. Transl. Med.* **2015**, *7*, ra138–ra303. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Wu, J.; Cui, Y.; Sun, X.; Cao, G.; Li, B.; Ikeda, D.M.; Kurian, A.W.; Li, R. Unsupervised Clustering of Quantitative Image Phenotypes Reveals Breast Cancer Subtypes with Distinct Prognoses and Molecular Pathways. *Clin. Cancer Res.* **2017**, *23*, 3334. [\[CrossRef\]](#)