Radiological tumor classification across imaging modality and histology

Jia Wu[#1,2,3,‡], **Chao Li**[#4,5], **Michael Gensheimer**[1], **Sukhmani Padda**[6], **Fumi Kato**[7], **Hiroki Shirato**[8], **Yiran Wei**[5], **Carola-Bibiane Schönlieb**[9], **Stephen John Price**[5], **David Jaffray**[2,10], **John Heymach**[3], **Joel W Neal**[6], **Billy W Loo Jr**[1], **Heather Wakelee**[6], **Maximilian Diehn**[1], **Ruijiang Li**[1,‡]

[1.]Department of Radiation Oncology, Stanford University School of Medicine, Palo Alto, CA, USA

[2.]Department of Imaging Physics, MD Anderson Cancer Center, Houston, TX, USA

[3.]Department of Thoracic and Head and Neck Medical Oncology, MD Anderson Cancer Center, Houston, TX, USA

[4.]The Centre for Mathematical Imaging in Healthcare, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, UK

[5.]Cambridge Brain Tumor Imaging Laboratory, Division of Neurosurgery, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

[6.]Department of Medicine, Division of Oncology, Stanford University School of Medicine, Stanford, CA, USA

[7.]Department of Diagnostic and Interventional Radiology, Hokkaido University Hospital, Sapporo, Hokkaido, Japan

[8.]Global Center for Biomedical Science and Engineering, Faculty of Medicine, Hokkaido University, Sapporo, Hokkaido, Japan

[9.]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, UK

[10.]Office of the Chief Technology and Digital Officer, MD Anderson Cancer Center, Houston, TX, USA

[#] These authors contributed equally to this work.

Abstract

[‡]**Corresponding Authors** Jia Wu, PhD., Department of Imaging Physics, Department of Thoracic/Head & Neck Medical Oncology, MD Anderson Cancer Center, 1400 Pressler St. - Unit 1472, Houston, Texas 77030, JWu11@mdanderson.org Phone: 713-563-2719, Ruijiang Li, PhD., Department of Radiation Oncology, Stanford University School of Medicine, 1070 Arastradero Rd. Palo Alto, CA 94304, rli2@stanford.edu; Tel.: 650-724-5382.

Radiomics refers to the high-throughput extraction of quantitative features from radiological scans and is widely used to search for imaging biomarkers for prediction of clinical outcomes. Current radiomic signatures suffer from limited reproducibility and generalizability, because most features are dependent on imaging modality and tumor histology, making them sensitive to variations in scan protocol. Here, we propose novel radiological features that are specially designed to ensure compatibility across diverse tissues and imaging contrast. These features provide systematic characterization of tumor morphology and spatial heterogeneity. In an international multi-institution study of 1,682 patients, we discover and validate four unifying imaging subtypes across three malignancies and two major imaging modalities. These tumor subtypes demonstrate distinct molecular characteristics and prognoses after conventional therapies. In advanced lung cancer treated with immunotherapy, one subtype is associated with improved survival and increased tumor-infiltrating lymphocytes compared with the others. Deep learning enables automatic tumor segmentation and reproducible subtype identification, which can facilitate practical implementation. The unifying radiological tumor classification may inform prognosis and treatment response for precision medicine.

Introduction

Radiological imaging is an integral component of cancer care and is used for screening, diagnosis, and staging, as well as for evaluation of treatment response and surveillance for disease relapse. Beyond its routine clinical applications, imaging can also provide rich information about tumor phenotypes, which are fundamentally governed by the underlying biological processes of the malignancy[1]. This is achieved by the high-throughput extraction of quantitative image features from standard-of-care radiological scans. This approach known as radiomics has been extensively applied in various cancer types with the goal of predicting therapy response and outcomes[2]. In addition, specific radiomic features have been linked to genetic and molecular characteristics of the tumor and its microenvironment[3–8].

Currently, radiomics analysis in any given study is focused on one imaging modality and one cancer type only. This is necessary because most radiomics features (e.g., texture) are highly sensitive to variations in image intensity, making them modality-dependent and histology-specific[9, 10]. In clinical practice, each disease has a preferred imaging modality that maximizes tissue contrast, for instance, computed tomography (CT) in lung cancer and magnetic resonance imaging (MRI) in breast cancer. As a result, it has been challenging to apply radiomics signatures identified in a cohort with a given disease and imaging modality to other settings, which limits their reproducibility and generalizability[11–13].

The Cancer Genome Atlas (TCGA) consortium recently performed an integrated molecular analysis of over 10,000 tumors in 33 cancer types[14]. With their greater statistical power, such pan-cancer studies may help identify commonly conserved patterns and unifying biological themes across cancers. Microsatellite instability (MSI), neurotrophin receptor tyrosine kinase (NTRK) gene fusions, and tumor mutational burden (TMB) are prime examples of tissue-agnostic biomarkers that are used to select patients for specific treatment regardless of tumor histology[15].

Given the availability of standardized genomic and transcriptomic data, current pan-cancer studies have mainly focused on molecular aspects of the tumor. In the context of radiological imaging, however, pan-cancer studies are notably lacking, significantly hampered by lack of standardization and diverse tissue contrast in different modalities. Compared with the conventional radiomic approach focusing on one cancer type and one imaging modality, a cross-histology and cross-modality strategy may lead to the discovery of unifying imaging phenotypes that are conserved across multiple cancer types and imaging modalities, which will have a broad impact on a larger population of patients.

In this study, we proposed novel radiological features to systematically characterize tumor morphology and spatial heterogeneity. These features were meticulously designed to ensure that they are comparable across diverse tissues and imaging contrast. Based on multi-institutional cohorts that span three cancer types and two major modalities, we identified and validated four unifying imaging subtypes that are associated with distinct oncogenic processes and prognoses after conventional therapies. We further demonstrated their potential predictive value in patients treated with immunotherapy.

Results

Overview of study design

We sought to define a radiological tumor classification scheme that is broadly applicable across cancer types and imaging modalities (Fig. 1a). To achieve this goal, we conducted an international multicenter study from 12 independent cohorts with a total of 1,682 cancer patients recruited from the US, Europe, and Japan (Supplementary Table 1). Our analysis included three cancer types (i.e., lung, breast, and brain malignancies); and two imaging modalities (i.e., CT and MRI). All patients had a pathologically confirmed diagnosis of primary malignancy, and had received standard therapies including surgery, radiation, chemotherapy, and/or hormonal therapy. We collected their pre-treatment radiological scans and, if available, tumor molecular profiles, clinicopathological and outcome data (Supplementary Tables 2–4). To rigorously assess reproducibility of our findings, we divided the 12 primary cohorts into discovery and independent validation sets as stratified by cancer types and imaging modalities. Further, we assessed the clinical relevance of the imaging subtypes in a separate cohort of 102 advanced lung cancer patients treated with anti-PD-1/PD-L1 immune checkpoint blockade (Supplementary Table 5).

Radiological features applicable across histology and imaging modality

We proposed two broad categories of quantitative features to characterize radiological phenotypes: tumor morphology and spatial heterogeneity (Fig. 1b–d). These features are specifically designed to account for diversity in tissue origin and imaging contrast. For morphology, we transformed the 3D tumor shape into a structured sequence of coefficients through spherical harmonic decomposition (Extended Data Fig. 1a). These coefficients provide an unbiased, optimal representation of the original tumor shape in the spatial frequency domain and can be used to fully reconstruct any tumor shape via a unique one-to-one mapping (Extended Data Fig. 1b–c). For spatial heterogeneity, we analyzed 3 distinct regions of interest including *tumor core*, *intratumoral invasive margin*, and *peritumoral*

parenchymal margin. Based on these non-overlapping regions, we defined 20 features to describe multiregional intensity variations through pair-wise comparison of normalized probability distribution functions. In total, 313 image features were extracted for each tumor.

We first confirmed that these image features were distributed similarly and thus comparable across imaging modality and cancer type (Extended Data Fig. 2). To test the reproducibility of extracted features against variation in tumor delineation, we randomly perturbed the original tumor contour and observed an overall high consistency with intraclass correlation coefficient of 0.83, indicating that our proposed features are robust to variation in tumor segmentation (Supplementary Fig. 1).

Efficient encoding of radiological features using machine learning

We split the patients in the 12 primary cohorts into discovery set and validation set stratified by cancer types, where the discovery set contains 601 lung cancer, 269 breast cancer, and 136 GBM, and the validation set contains 360 lung cancer, 185 breast cancer, and 131 GBM. Given the large number of features, we trained an artificial neural network (ANN) called autoencoder in discovery cohort to learn efficient representations of the original features and mitigate the curse of dimensionality (Extended Data Fig. 3a). The trained autoencoder was tested in the validation set (Extended Data Fig. 3b). Instead of linear PCA, we used autoencoder for feature representation due to its capability of modelling complex non-linear patterns. After training, the autoencoder substantially reduced the dimension of features to 10 from 313, while at the same time maximally preserving information in the original data (Extended Data Fig. 3b–c). The feature dimension reduction has the dual advantages of effectively removing redundancy in the features and improving stability and efficiency of subsequent clustering analysis.

Discovery and validation of tumor subtypes across histology and imaging modality

In the discovery set (n = 1006), based on the consensus clustering of radiological tumor phenotypes encoded by 10 autoencoder features, we determined the optimal cluster number to be four (Fig. 2a, Supplementary Fig. 2a), which maximized consensus within clusters while minimizing ambiguity in cluster assignments. Next, we independently applied the same consensus clustering analysis and trained autoencoder in the validation set (n = 676) and also identified 4 clusters of patients (Fig. 2b, Supplementary Fig. 2b). To measure the reproducibility of clusters across the discovery and validation sets, we computed the in-group proportion (IGP) statistics, which showed high consistency for 4 clusters, with IGP values of 92.4% ($P < $ 1e-10), 91.1% ($P = 0.002$), 83.2% ($P = 0.008$), 78.0% ($P = 0.021$), respectively.

We then split the patients based on imaging modalities, i.e., CT vs. MRI. Based on the discovery set of CT features in 961 lung cancer patients, we again identified 4 clusters (Fig. 2c, Supplementary Fig. 2c). Similarly, four clusters were identified in the validation set based on MRI features in 454 breast cancer and 267 GBM patients (Fig. 2d, Supplementary Fig. 2d). These clusters were found to be highly consistent across imaging modality, with IGP values of 81.0% ($P = 0.032$), 80.0% ($P = 0.041$), 91.8% ($P = 0.014$), and 91.2% ($P = 0.008$), respectively.

After confirming the consistency of patient clusters across imaging modalities and cancer types, we refined the subtypes by using all patients in the primary cohorts (Fig. 2e, Supplementary Fig. 2e). Overall, this resulted in 580 patients (34.5%) in cluster 1, 647 patients (38.5%) in cluster 2, 272 patients (16.2%) in cluster 3, and 183 patients (10.9%) in cluster 4. All four clusters were represented in each of the three cancer types (Extended Data Fig. 4a). The distribution of clusters was largely independent of cancer type (Cramér's V = 0.21, P = 0.073). We further evaluated the accuracy of the clustering results (see details in Methods). Overall, the cluster purity scores were 91–97%, confirming validity of the clustering results (Fig. 2f). Within each cancer type and imaging modality, the clusters remained highly reproducible. These histology and modality-independent clusters are thereafter named imaging subtypes 1 through 4.

Imaging characteristics distinguishing the tumor subtypes

To better understand specifically which factors contribute to the distinction among the subtypes, we performed differential analysis by using the original 313 imaging features while controlling for multiple testing (Extended Data Fig. 5). We identified four main categories of image features associated with the subtypes, which measure tumor volume, shape symmetry, shape regularity, and regional variation (Fig. 3a–b). Based on this analysis, we summarized the radiological interpretations of four subtypes in Fig. 3c. Specifically, Subtype 1 mainly consisted of small tumors with large variations across tumor-parenchyma interface (i.e., distinct, sharp margin); Subtype 2 consisted of intermediate sized tumors with moderately well-defined margin; Subtype 3 consisted of large tumors with ill-defined, blurred margin; Subtype 4 was characterized as large tumor with moderately well-defined margin. For tumor morphology, subtypes 1 to 3 had similarly smooth and regular shape, while Subtype 4 had the highest complexity with rugged and irregular shape. Figure 3d shows a graphical representation of the 4 imaging subtypes in the feature space.

Clinical validation for the prognostic significance of the imaging subtypes

We tested the prognostic relevance of the imaging subtypes and observed significant differences in patient prognosis within each cancer type (Fig. 4a–c). In NSCLC, subtype 1 was associated with the best survival, subtypes 2 and 3 with intermediate prognoses, and subtype 4 with the worst prognosis. Consistently, in breast cancer and GBM, patients in subtype 4 also had the worst prognosis. In multivariate analysis, subtype 4 remained an independent predictor for poor survival after adjusting for stage, tumor volume, clinicopathological factors, and specific molecular features (Fig. 4d–f, Supplementary Table 6–7).

We further evaluated the prognostic significance of the imaging subtypes by focusing on patients in clinically relevant subgroups. Specifically, we confirmed that their prognostic value was largely independent of disease stage and primary treatment in NSCLC. The distribution of subtypes for clinical stage in NSCLC is shown in Extended Data Fig. 4b. Moreover, the imaging subtypes remained prognostic within early (stage I/II) as well as locally advanced (stage III) NSCLC (Extended Data Fig. 6a–b). One notable exception is subtype 4, which was associated with a worse prognosis compared with subtype 3 for patients treated with surgery (HR=1.82, 95% CI 0.89–3.75, P=0.064). On the other hand,

subtype 4 appeared to have an improved prognosis that is similar to subtype 3 for patients treated with radiotherapy (HR=0.94, 95% CI 0.61–1.46, P=0.8), (Extended Data Fig. 6c–d). Fewer patients in subtypes 3 and 4 received surgery compared with radiotherapy (n = 67 vs 160). We tested the interaction between therapeutic regimen (surgery vs radiotherapy) and subtype 3 vs 4 and found a possible interaction effect (P = 0.063). Moreover, we performed subgroup analysis by stratifying tumors according to driver gene alterations in NSCLC, such as EGFR and ALK (Extended Data Fig. 7a–c).

Similarly, we tested the imaging subtypes in different histological subtypes of breast cancer and confirmed prognostic significance especially in HER2-positive patients (Extended Data Fig. 7d–f). For GBM, we tested the survival stratification by imaging subtypes in different patient subgroups defined by MGMT methylation or IDH1 mutation status (Extended Data Fig. 7g–i).

To further demonstrate the advantage of using these cross-modality and cross-histology image features, we compared the performance for prognostic prediction using the conventional radiomic approach. Specifically, we trained a radiomic model in a supervised manner to predict survival of lung cancer patients (Extended Data Fig. 8a–b). We observed that the radiomics risk score was highly correlated with tumor size (Pearson correlation coefficient = 0.92, Extended Data Fig. 8c), and the radiomics model had a C-index of 0.60 (95% CI: 0.56–0.64) in the validation cohort. In comparison, our proposed imaging subtypes were independent of tumor volume and achieved a C-index of 0.67 (95% CI: 0.63–0.72) in validation, which is significantly better than radiomics model (P < 0.001, Extended Data Fig. 8d).

Biological validation for the molecular basis of the imaging subtypes

We performed gene set enrichment analysis to identify molecular pathways associated with the imaging subtypes. In NSCLC, for subtype 1 with the most favorable prognosis, the majority of cancer hallmark pathways, including proliferation, angiogenesis, and hypoxia, were significantly downregulated compared with other subtypes, (Extended Data Fig. 9a, Supplementary Table 9). By contrast, many cancer hallmarks including glycolysis and metastasis-related pathways were upregulated in subtype 4, suggesting a more aggressive phenotype consistent with the poor survival in these patients (Extended Data Fig. 9b, Supplementary Table 10). Of note, we also observed an increased expression of genes related to radiation response such as cell cycle, apoptosis, and DNA repair, consistent with their increased survival when treated with radiotherapy.

Similar enrichment analyses were also performed separately for breast cancer and GBM. However, we did not observe any pathways that reached the predefined statistical significance level at FDR < 0.05, likely owing to limited statistical power given a small number of samples (Extended Data Fig. 9c–f). We further investigated the distribution of established molecular subtypes among the different imaging subtypes (Extended Data Fig. 4c–d), and they were independent of each other, with Cramér's V = 0.18 (P = 0.539) for breast cancer and Cramér's V = 0.16 (P = 0.490) for GBM.

Imaging subtypes stratify survival in patients treated with immunotherapy

The imaging subtypes were discovered and validated in 12 primary cohorts of patients treated with conventional therapies. We sought to further evaluate their clinical relevance in the immunotherapy setting in a completely independent cohort. To this end, we collected data from an institutional cohort of 102 advanced NSCLC patients who were treated with immune checkpoint blockade. Given the extracted CT features of primary tumors, we trained a multiclass classification model using XGBoost (Supplementary Methods) to predict the imaging subtype of each new patient. The majority of these patients with advanced NSCLC were classified as either subtype 3 (37%) or subtype 4 (50%), as shown in Fig. 5a, and both subtypes had large tumors. This is consistent with the fact that advanced-stage tumors tend to be larger than early-stage tumors. We then confirmed that the distributions of characteristic imaging features followed the same patterns in both the primary and immunotherapy cohorts (Fig. 5b). In particular, tumors in subtype 3 had regular shape and ill-defined, blurred margin, whereas those in subtype 4 had irregular shape and moderately well-defined margin.

There was no statistically significant difference in survival between patients in subtypes 1&2 and subtype 4 (Extended Data Fig. 10). Strikingly, we observed that patients in subtype 4 had significantly better survival compared with subtype 3 (HR = 0.46, 95% CI: 0.23–0.93, P = 0.034), as shown in Fig. 5c. This is in stark contrast with the results in our primary cohorts treated with conventional therapies, where patients in subtype 4 had a significantly worse prognosis. We further compared the immune cell composition in the tumor microenvironment between subtypes 3 versus 4. Tumors in subtype 4 had a higher infiltration of several adaptive immune cell populations, including activated CD56^{dim} natural killer (NK) cells, cytotoxic CD8 T cells, CD4 T helper cells, and $\gamma\delta$ T cells (Fig 5d, Supplementary Table 11).

Deep learning enables automated tumor segmentation and reproducible identification of imaging subtypes

Our previous analysis of the imaging subtypes requires the tumor contour, which involves manual delineation and is subject to inter- and intra-rater variability. To overcome this issue and facilitate practical implementation, we trained a deep learning model based on 2D U-Net (Fig. 6a) to perform automated tumor segmentation. The U-Net model performed well at segmenting tumors in NSCLC, with average DICEs of 0.90, 0.88, and 0.84 for training, validation, and testing, respectively (Fig. 6b). This model performed especially well for large tumors (>10 cc), with DICE > 0.9, and for tumors not attached to mediastinum, with DICE > 0.85 (Fig. 6c–d). On the other hand, the U-Net model did not perform well in breast cancer, with DICEs of 0.73, 0.72, and 0.69 for training, validation, and testing, respectively (Supplementary Fig. 3). For GBM, tumor segmentation was performed with a pre-trained deep learning model.

Finally, we applied the automatically generated tumor segmentations of lung patients and repeated the entire pipeline for imaging subtype discovery (Fig. 6e). The clustering results were highly reproducible, with the cluster assignment for the majority (83.5%) of patients

remaining unchanged based on manual versus deep learning generated segmentations (Fig. 6f).

Discussion

In this international multi-cohort study of 1,682 patients, we discovered and independently validated novel tumor subtypes that are broadly applicable across major imaging modalities and three cancer types. These subtypes demonstrate distinct radiological and molecular features, as well as survival outcomes after conventional therapies. Moreover, their prognostic value was independent of established clinical risk factors including tumor volume. Importantly, we showed that specific imaging subtypes are associated with differential outcomes after immunotherapies. Finally, we demonstrated that deep learning can be used to automate tumor segmentation, which will help standardize subtype identification and facilitate its implementation in clinical practice.

Radiologic-based classification has several important advantages compared with histopathologic assessment or molecular tumor profiling. First, imaging evaluation is performed for the entire 3D tumor in situ, which overcomes the sampling bias of a small biopsy due to intratumoral spatial heterogeneity. Second, imaging provides a non-invasive means for evaluation, which would allow longitudinal monitoring and follow-up of disease. Because radiomic features provide a representation of tumor phenotypes from a radiological perspective, the proposed radiologic-based tumor classification could complement current clinical and molecular classification. We envision that integration of information from different domains may further enhance prediction of treatment response and prognosis for precision oncology[16, 17].

It is important to emphasize that the purpose of our work is not to develop a predictive model for a *specific* clinical endpoint (such as prognosis) in a *particular* cancer as is done in the traditional radiomics approach with supervised machine learning. Rather, our study is aimed at the discovery of unifying radiological phenotypes across different imaging modalities and cancer types. By identifying these converging radiographic hallmarks, we further revealed biological insights and established their connection with prognosis and response to different cancer treatments including surgery, radiation and chemotherapy, as well as immunotherapy.

Previous works on radiomics have been limited to one cancer type and one imaging modality in any given study. Consequently, it is difficult to generalize the findings that are often based on relatively small datasets[13]. Here, we aimed to discover common radiological patterns with biological and clinical relevance across histology and modality. To achieve this goal, we meticulously designed image features that are robust to tissue contrast. This study provides a conceptual framework that will allow aggregation of datasets with disparate modalities and cancer types, similar to integration of molecular data in pan-cancer studies by TCGA.

Despite extensive investigations in radiomics, the progress toward clinical translation has been slower than anticipated[12, 13]. One major reason is a lack of reproducibility and rigorous

validation, given the exploratory nature and inclusion of small numbers of patients in many studies[12, 18]. Textural features, which usually represent the overwhelming majority of features in a typical radiomics study, have been shown to be highly sensitive to imaging protocol and technical factors[9]. This has precluded meaningful comparison of existing radiomics signatures across studies with different modalities and various cancer types[11].

To overcome these issues, we focused on two broad categories of radiological features (i.e., tumor morphology and spatial heterogeneity) and made special efforts to ensure their compatibility across imaging modalities. Geometric features are computed based on a binary image of the tumor contour, which are insensitive to tissue contrast and can be readily comparable across different imaging modalities and cancer types[19]. Specifically, we applied the spherical harmonic decomposition method to approximate the 3D tumor shape with a series of orthonormal basis functions defined on the surface of the sphere. Similar to Fourier analysis of time series, this allows us to interrogate in an unbiased manner the geometric patterns with an ordered sequence of coefficients in the frequency domain. Previous morphological features only focus on the low-frequency component (Extended Data Fig. 1c, degree 1) and disregard rich information captured at higher spatial frequencies (degrees 2 to 15).

While texture features are widely used as a measure of tumor heterogeneity, current approaches typically do not distinguish different regions within the tumor; most are focused on the primary tumor, ignoring the surrounding parenchyma. There is emerging evidence supporting the clinical relevance of these different regions[20, 21]. Here, we measure tumor heterogeneity by performing spatially explicit analysis of intratumoral and peritumoral regions as well as the invasive margin. By making pair-wise comparison of normalized density functions and using each region as an internal control, we generated features that are robust to variability in tissue contrast. It is worth noting that our tumor subregions are defined from a 'geometric' perspective, which is simple to implement and is generally applicable across cancer types and imaging modalities. On the other hand, physiological tumor subregions or habitats such as hypoxia and hypermetabolic activity can provide more meaningful biological information, and they do not necessarily follow this simple geometric paradigm[22, 23]. One caveat is that these physiological tumor subregions are likely cancer type-specific and imaging modality-dependent, and reliable identification of these subregions requires sophisticated algorithms, such as habitat imaging[24, 25].

Here, we showed that the imaging subtypes are associated with distinct survival outcomes in three cancer types after conventional therapies. Further, the prognostic value of our imaging subtype was independent of tumor volume and other clinical factors. A key limitation of previous radiomic studies is that some signatures may be correlated with tumor burden, which would diminish their clinical relevance[26]. Similarly, we also observed that radiomics model was highly correlated with tumor size (Extended Data Fig. 8c). This underscores the importance that any useful biomarker must provide additional value beyond established risk factors.

The new imaging subtypes may have therapeutic implications. In particular, we found that subtype 4 was associated with improved survival in NSCLC patients treated with immune

checkpoint blockade. Consistently, we found a higher infiltration of cytotoxic NK and T lymphocytes in the tumor microenvironment of subtype 4. These results suggest that subtype 4 tumors are associated with a pre-existing antitumor immunity and thus may preferentially respond to immunotherapy[27]. Interestingly, a recent study developed a CT-based radiomics signature specifically for estimating tumor-infiltrating CD8 T cells, which was correlated to survival after immunotherapy[4]. These findings are based on retrospective analysis, and will require prospective validation in future randomized trials.

For clinical implementation, different imaging modalities used in different malignancies may introduce some practical challenges. For instance, in lung cancer imaging, normal breathing can induce some degree of tumor motion depending on its anatomic location, which may cause blurring artifacts in the CT image. Methods for dynamic lung modeling and tumor tracking may be useful[28]. Modern multi-slice CT scanners allow fast imaging with a breath-hold protocol and can largely mitigate this issue. On the other hand, MRI is subject to its own source of variability due to the use of various sequence protocols and parameters. Further development in quantitative imaging with standardized acquisition should improve image quality and reproducibility of radiomic biomarkers[29]. In breast cancer, MRI may be acquired after tumor biopsy, which can introduce imaging artifacts and affect the calculation of certain features related to spatial variation.

One limitation of the radiogenomic analysis is that gene expression profiling was performed for the bulk tumor from surgical specimens. Single-cell gene expression analysis can provide much refined details and offer a deeper insight about tumor heterogeneity[30]. In future work, the radiogenomic association findings should be confirmed at the protein level using immunohistochemistry or immunofluorescence assays. Further, it will be important to establish a 'mechanistic' link between the imaging subtypes and their biological underpinnings, which will require gene knock-in/knock-out experiments using in vivo animal studies.

Deep learning such as convolutional neural networks has emerged as a powerful technique for medical image analysis and achieved promising performance in various clinical applications[31–33]. Future development of advanced deep learning techniques including physics-reinforced or physics-award algorithms[34–36] that can lead to further improvement for more reliable automated tumor segmentation[37], which will allow consistent identification of imaging subtypes defined here.

Finally, we note that while finding unifying imaging phenotypes across cancers are useful as demonstrated here, there certainly exist modality- and histology-specific features that are also important in determining disease outcomes. For instance, the edematous and diffusive growth pattern as visualized by specific MRI sequences is unique to GBM[38]. In this study, we mainly investigated anatomical imaging with CT and MRI, two most widely used modalities in clinical oncology practice. Our result does not contradict the ongoing or future investigation of imaging characteristics that are specific to modality or cancer type. More advanced, specialized imaging technologies may allow direct visualization of the functional and molecular characteristics, which could provide complementary information about the disease.

In conclusion, we have proposed a radiological tumor classification system that is applicable across imaging modality and histology. These imaging subtypes are associated with distinct oncogenic and microenvironmental features as well as survival patterns. Future studies are needed to validate the potential of this system to identify patients who are likely to benefit from immunotherapy.

Methods

Study design and patient cohorts

For subtype discovery and validation, we included a total of 12 cohorts: 6 non-small cell lung cancer or NSCLC cohorts (n=961) with CT scans, 4 breast cancer cohorts (n=455) and 2 glioblastoma multiforme or GBM cohorts (n=266) with MRI scans. Among these, data for 3 cohorts (one for each cancer type) were retrospectively collected from each participating center and the remaining 9 cohorts are publicly available. Follow-up and outcome data were available for 1,289 patients. We evaluated prognostic significance of the imaging subtypes in different clinical settings. Further, using tumor gene expression data available for 652 patients, we investigated molecular correlates of the imaging subtypes. Finally, we collected data from an independent cohort of 102 patients with advanced lung cancer treated with anti-PD-1/PD-L1 immune checkpoint blockade and assessed clinical relevance of imaging subtypes in the immunotherapy setting. Detailed clinical characteristics for the discovery cohorts and the immunotherapy validation cohort are summarized in Supplementary Tables 2–5 (see also Supplementary Methods).

Scan acquisition and image processing

We collected pre-treatment scans for the most commonly used imaging modality for each cancer, *i.e.*, CT for NSCLC, MRI for breast and GBM. Given the heterogeneous scan protocols, we applied a series of image processing algorithms to harmonize the image data and facilitate robust feature extraction. The primary tumor was manually delineated in all slices by experienced physicians and the surrounding lung, brain, and breast parenchyma was automatically segmented. The details about scan protocols, image processing, and segmentation are presented in Supplementary Methods.

Radiological features

We defined two broad categories of quantitative features, i.e., tumor morphology and spatial heterogeneity. For morphology, we transformed the 3D tumor shape into a structured sequence of coefficients through spherical harmonic decomposition (Extended Data Fig. 1a). Special efforts are made to ensure that the shape decompositions are invariant to scale (size), translation, and rotation. Based on these coefficients, we proposed additional second-order features to summarize the shape complexity, including shape irregularity that measures the boundary smoothness as well as shape symmetry that measures directionally dependent tumor expansion. We extracted five features to characterize tumor burden, resulting in 293 shape descriptors ($\vec{f}_S$). For spatial heterogeneity, we analyzed 3 distinct regions of interest given the tumor contour and its surrounding parenchyma, designated as tumor core, intratumoral invasive margin, and parenchymal (peritumoral) margin. Based on these

nonoverlapping regions, we defined 20 features ($\vec{f}_R$) to describe multiregional intensity variations through pair-wise comparison of normalized probability distribution functions. The details of the image features and their interpretations are explained in Supplementary Methods and Table 12. In total, 313 quantitative image features were extracted for each tumor.

In order to evaluate the sensitivity of feature extraction with respect to variation in tumor segmentation, we generated a new set of tumor segmentation through random perturbation of the original tumor contour by applying elastic deformation in 100 randomly selected patients via Matlab function *affine3d*. We then repeated the pipeline for image feature extraction and computed the intraclass correlation coefficient between the two sets of extracted image features.

We trained and validated an autoencoder, a type of artificial neural network used for dimensionality reduction and representation learning, to efficiently encode the original 313 image features in a low-dimensional space (Extended Data Fig. 3). We chose autoencoder rather than linear PCA to account for complex non-linear relationship among high dimensional features. The dimensionality reduction will avoid undue influence of redundant features and also reduce noise. Details are described in Supplementary Methods.

Discovery and validation of imaging subtypes

We identified intrinsic imaging subtypes by applying unsupervised consensus clustering to study patients from 12 multi-center cohorts (see details in Supplementary Methods). The patients were divided into discovery and validation sets based on cancer types and imaging modalities, to assess reproducibility of the clustering analysis and prevent information leakage. As there were multiple cohorts, the largest cohort within each cancer type was preselected as the validation set and the remaining ones were merged into the discovery set. The consensus clustering analysis was carried separately in the discovery and validation cohorts with identical configurations. In detail, the partition around medoids clustering algorithm[39] with the Pearson's correlation distance metric was used as the basis for clustering. We performed 10,000 bootstraps with 80% item resampling of the autoencoder features. The optimal number of clusters was determined by varying from 2 to 10 and finding one that produced the most stable consensus matrices and the most unambiguous cluster assignments across permuted runs. The in-group proportion (IGP) statistic[40] was applied to measure the similarity of imaging subtypes identified between the discovery and validation sets. IGP ranges from 0% to 100%, where higher value indicates more similarity between 2 data sets.

With confirmation of the highly consistent imaging subtypes across discovery and validation cohorts, we refined the final imaging subtypes pulling all patients in the 12 cohorts. To get a robust subtype definition, the reproducibility of clustering analysis was evaluated in repeated 5 fold cross-validation scheme, where patients were randomly separated into training and testing sets at a ratio of 4:1. The training set was used to build the clustering model, which in turn was used to predict patient labels in the hold-out testing set. Cluster purity[41] was used to evaluate the clustering robustness. Cluster purity ranges from 0% to 100%,

where higher value indicates more robustness of clustering results. Finally, we identified specific radiological features that were significantly associated with the newly identified imaging subtypes. Specifically, the single-sample GSEA (ssGSEA) algorithm[42] was applied to compute the enrichment scores, and the limma algorithm[43] was implemented to model the differential expression in imaging features.

Clinical validation for prognostic significance

For 1,289 patients with detailed follow-up information, we evaluated prognostic relevance of the imaging subtypes. Within each of the three cancer types, we adjusted for established clinicopathological risk factors in the multivariate analysis. For NSCLC (n=701), we adjusted for age, gender, stage, tumor volume; for breast cancer (n=226), we included age, ER, PR, HER2 status, and tumor volume; for GBM (n=260), we considered age, gender, MGMT methylation, IDH1 mutation status, and tumor volume. We further evaluated prognostic significance of the imaging subtypes in clinically relevant subgroups, such as primary treatment (surgery vs. radiotherapy) in NSCLC and established molecular subtypes in breast cancer and GBM.

Biological validation to identify molecular correlates

We performed gene set enrichment analysis (GSEA) to identify molecular pathways associated with each of the imaging subtypes. This was done in a subset of patients for whom both imaging and gene expression data are available, i.e., NSCLC (n=274), breast cancer (n=254), and GBM (n=124). Specifically, the single-sample GSEA (ssGSEA) algorithm[42] was applied to compute the enrichment scores of 50 cancer hallmark pathways curated from MSigDB. Moreover, ssGSEA is a powerful way to mitigate batch effects and overcome discrepancies between different technologies (RNA-seq vs. microarray). The limma algorithm[43] was implemented to model the differential expression pathways.

Clinical evaluation in lung cancer immunotherapy

We analyzed an independent cohort of 102 advanced NSCLC patients treated with anti-PD-1 or anti-PD-L1 immune checkpoint blockade, and assessed clinical outcomes in relation to the previously identified imaging subtypes. For imaging feature extraction, we focused on the lesion with the largest size. Since this cohort was not included in the imaging subtype discovery and validation sets, we built an ensemble learning classifier (i.e., XGBoost, Supplementary Methods) to predict the imaging subtype label of each new patient given the radiological features bypassing the procedure of dimension reduction and clustering. Since the immune cell composition in tumor microenvironment may influence response to immunotherapy, we assessed the enrichment of tumor-infiltrating immune cells for each individual imaging subtype. In particular, we took the gene expression for 274 primary NSCLC tumor samples and then estimated the abundance of 16 immune cell populations by using a previously curated list of genes (Supplementary Table 13).

Evaluation of subtype reproducibility with deep learning of automated tumor segmentation

We trained a deep convolutional neural network (i.e., U-Net[44]) to fully automate the tumor segmentation process (see details in Supplementary Methods). The main goal here

is to reduce the inter-observer variations and improve consistency in tumor segmentation by standardizing this process. Further, automated segmentation can facilitate clinical implementation of the tumor subtyping approach by saving time and annotation efforts. Sørensen–Dice coefficient was used to measure the quality of automated segmentation, with manual contours serving as the ground truth. The automated contours were used as input to repeat the entire computational analysis and assess reproducibility of the previously identified imaging subtypes.

Statistical analysis

Kaplan-Meier analysis and logrank tests were used to evaluate statistical significance of patient stratification by the imaging subtypes. Cox proportional hazard regression model was used to adjust for relevant clinicopathologic variables in multivariable analysis. For differential expression analysis, the Benjamini-Hochberg method was used to adjust for multiple statistical testing and control the false discovery rate (FDR). The Cramér's V statistic was used to assess correlation between imaging clusters and cancer types. All statistical tests were two-sided, with a p-value less than 0.05 or FDR less than 0.05 considered statistically significant. All statistical analyses were performed in R.

Data availability

The data is available within the Article or the Supplementary Information. The imaging data for 9 out of totally 13 cohorts used in this study are publicly available through TCIA website (https://www.cancerimagingarchive.net/) as described in the Supplementary Information. The imaging data for the breast cancer cohort from Hokkaido University, Japan is publicly available at https://drive.google.com/drive/folders/1AsI-bvUWwdmwMd7SHXzJttUsKqmImAGz?usp=sharing

The imaging data for Stanford Lung Cancer, Lung Cancer Immunotherapy, and Cambridge GBM cohorts are not publicly available because they contain sensitive information that may comprise patient privacy as well as the ethical constriction or the regulation policy of local institutions. These data will be made available to individuals who contact the corresponding authors with a reasonable request, e.g., for non-commercial, research purposes. The gene expression data and mutational data of TCGA samples are publicly available in the Genomic Data Commons (https://gdc.cancer.gov/). The gene expression data for the other cohorts are available from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/; [accession number: GSE22226, GSE103584, GSE58661]).

Code availability

For the spherical harmonic decomposition, we used the SPHARM-MAT software (http://www.iu.edu/~spharm/). For autoencoder, XGboost, consensus clustering, we used R software (version 3.5.3, R Foundation for Statistical Computing, Vienna, Austria); package autoencoder (version 1.1), XGboost (version 1.1.1.1), ConsensusClusterPlus (version 1.52.0). The U-Net architecture is available at https://github.com/lyakaap/Kaggle-Carvana-3rd-place-solution Custom codes[45] are available at https://github.com/WuLabMDA/PanCancer

Extended Data

A

Input: 3D tumor contour, $t(x,y,z)$

Parameterize to spherical coordinate:
$\vec{t}(\theta,\phi)$,
$\theta \in [0,\pi], \phi \in [0,2\pi)$

Normalize to unit volume:
$\vec{f}(\theta,\phi)$,
$\theta \in [0,\pi], \phi \in [0,2\pi)$

Spherical Harmonic Expansion

$$\vec{f}(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \vec{c}_l^m Y_l^m(\theta,\phi)$$

- $\vec{c}_l^m$ is the spherical harmonic coefficients
- $Y_l^m(\theta,\phi)$ is the spherical harmonic basis
- l is degree, higher value means high-frequency (finer) details
- m is order, for any given degree, it captures different shape variation components with same level of frequency, where negative components are complex conjugate for its positive counterparts.

B

C

Figure 1.

Morphological characterization of tumors by spherical harmonic decomposition. a) Overall design of morphological analysis; b) Illustration of 3D spherical harmonic basis functions at different degrees and orders; c) Illustration of 3D tumors reconstructed by coefficients obtained from spherical harmonic decomposition. Each row represents a selected 3D tumor, which is reconstructed using decomposition results at 5 different degree levels. Here, lower degree captures more global patterns and higher degree corresponds to more detailed morphological patterns.
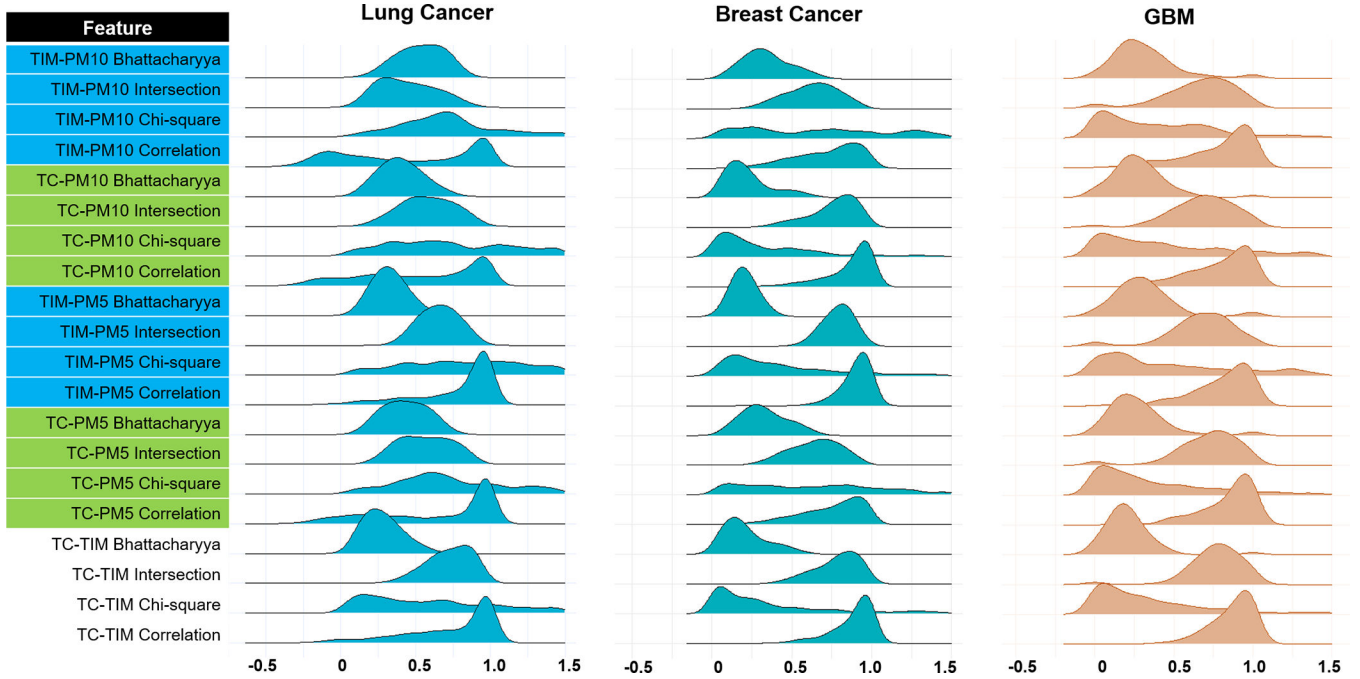
Figure 2.

The ridgeline plots present the distribution of 20 regional variation features in three different cancer types. Here, we investigate 2 tumor regions, tumor core (TC) and tumor invasive margin (TIM), plus 2 peritumor regions, parenchymal margin at 5 mm or 10 mm (PM5 or PM10). In total, 5 pair-wise regions are considered, namely, TC-TIM, TC-PM5, TC-PM10, TIM-PM5, TIM-PM10. Variation for each pair-wise region was quantified with four measures (chi-square, Bhattacharyya distance, correlation, intersection), yielding $5*4 = 20$ regional variation features. TC-PM5 and TC-PM10 related features are colored in green, while TIM-PM5 and TIM-PM10 related features are colored in blue.
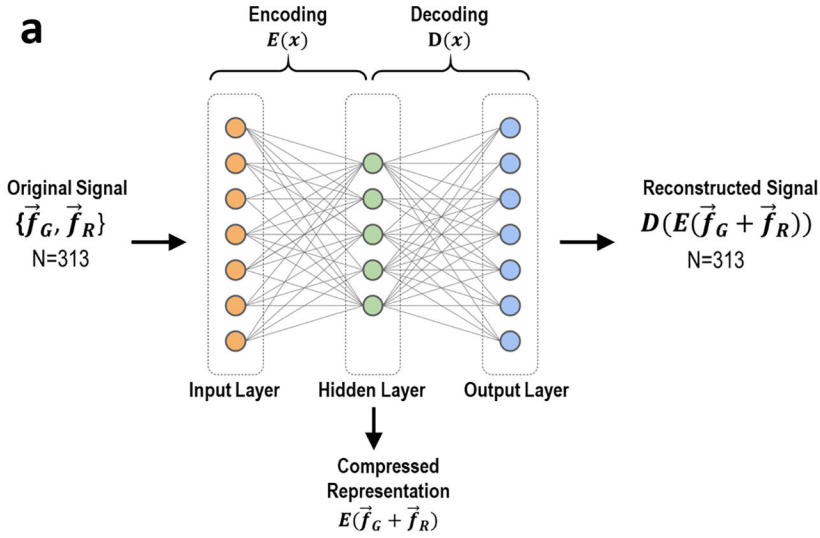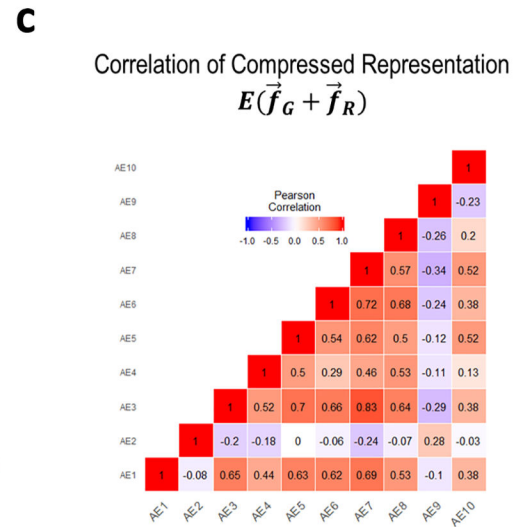
Figure 3.
Details of imaging feature dimension reduction via an autoencoder model. a) The structure of autoencoder used to learn a low-dimensional mapping of the original feature signals with detailed tuning hyperparameters; b) The optimal autoencoder loss curves in training and validation; c) Heatmap of pairwise correlations between 10 autoencoded features.
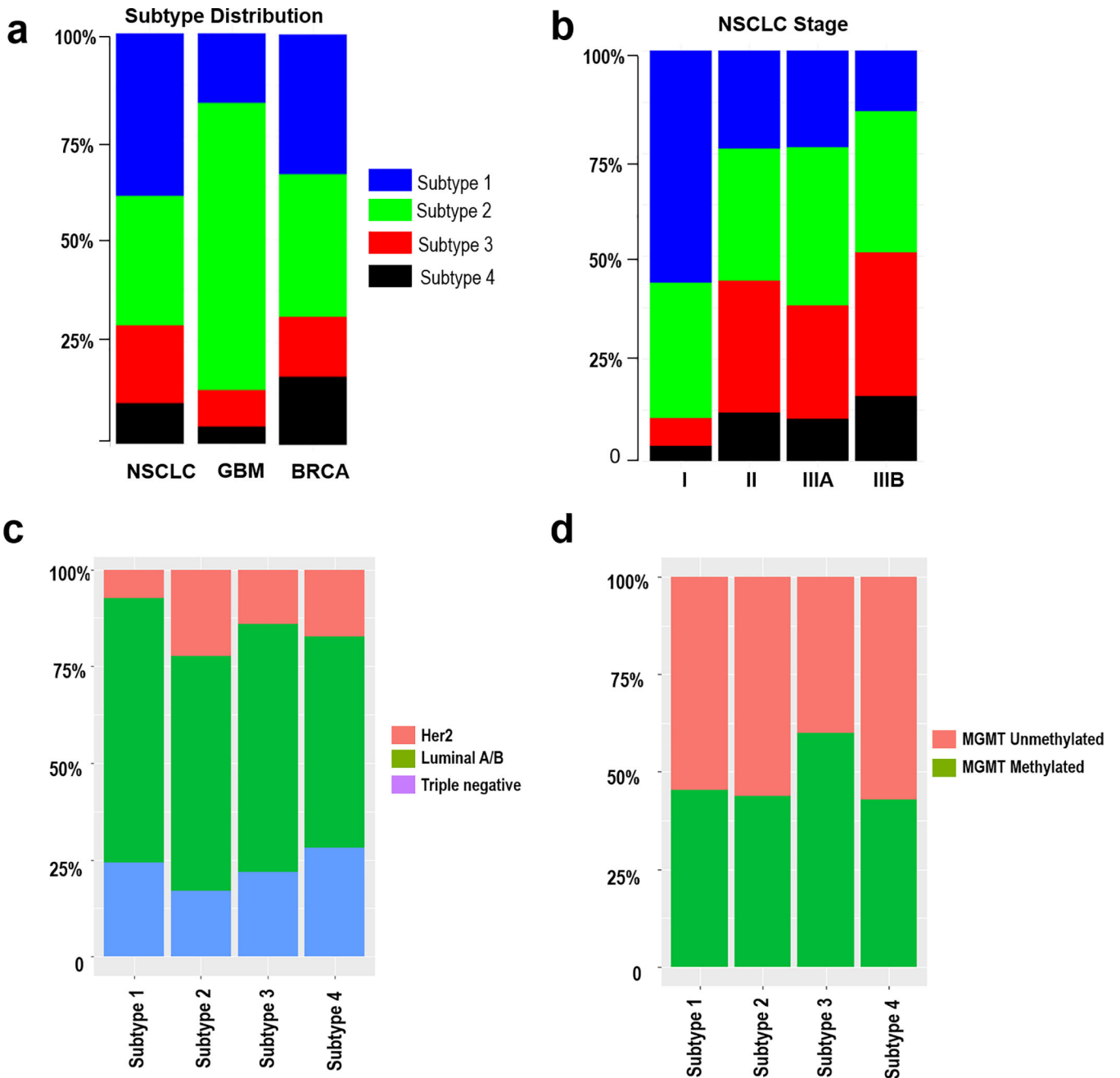
Figure 4.
Distribution of imaging clusters (subtypes) in different clinical groups. a) The distribution of all patients in four clusters (subtypes) across three cancer types; b) The distribution of lung cancer patients in four clusters (subtypes) across different clinical stage; The molecular subtype distribution in four imaging subtypes for c) breast cancer with luminal A/B, Her2+, and triple negative; d) GBM with different MGMT methylation status.
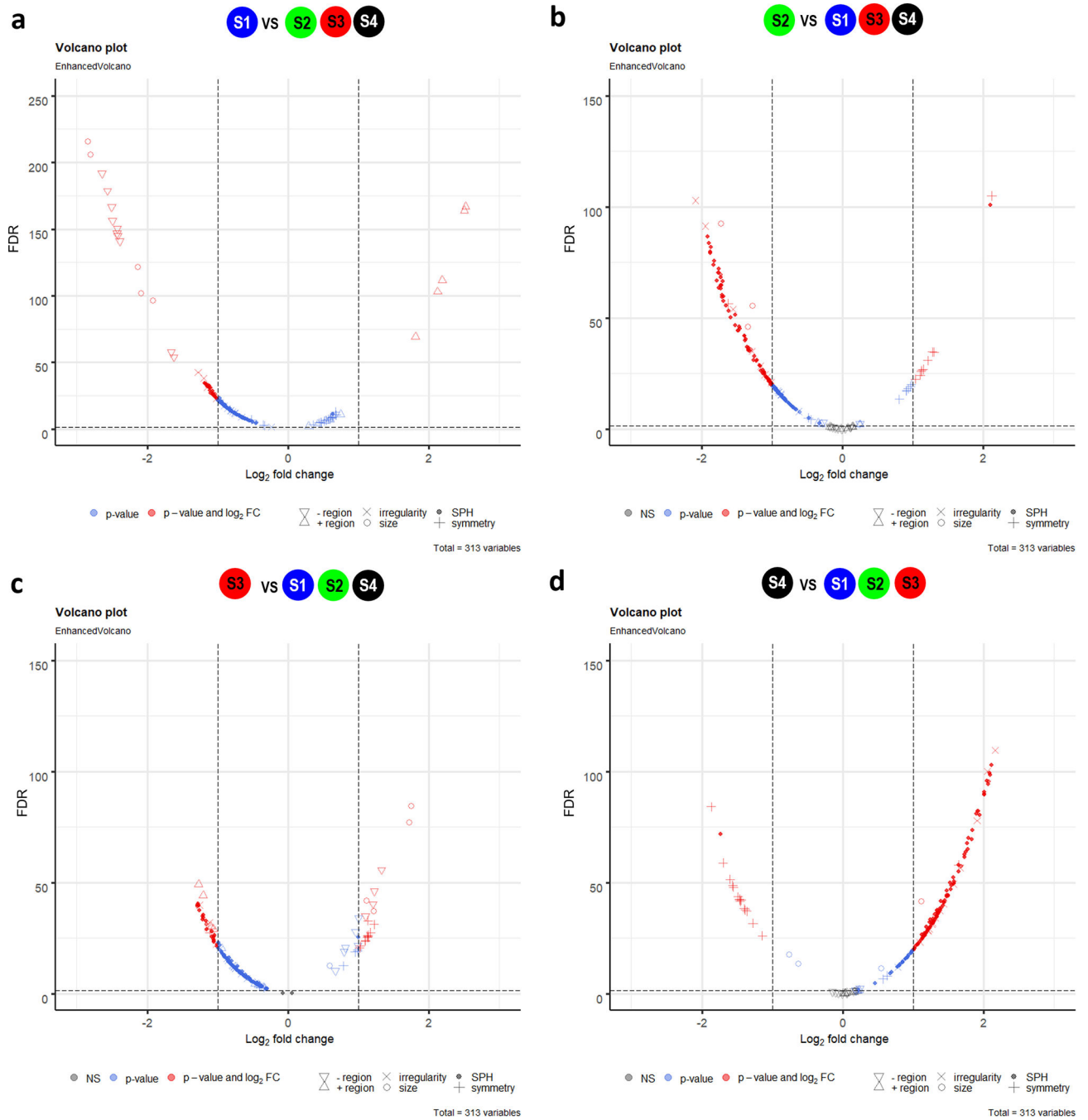
Figure 5.
Volcano plot of enrichment scores through single-sample Gene Set Enrichment Analysis (ssGSEA) of 313 proposed imaging features in all three cancer types. a) imaging subtype 1 versus rest, b) subtype 2 versus rest, c) subtype 3 versus rest, and d) subtype 4 versus rest. The data for all enrichment scores are plotted as log2 fold change versus the −log10 of the adjusted p-value. Thresholds are shown as dashed lines. Pathways deemed as significantly different (false discovery rate or FDR<0.05) are highlighted with different color schemes.

a

NSCLC Stage I/II (n=378)

logrank P = 3e-5

S1	186	154	108	82	52	32
S2	124	81	48	28	14	8
S3	47	26	16	12	9	6
S4	21	15	6	3	3	2

b

NSCLC Stage III (n=343)

logrank P = 0.003

S1	65	42	19	11	8	6
S2	125	74	25	13	8	4
S3	108	44	25	15	8	7
S4	45	17	5	2	2	1

c

NSCLC Surgery (n=277)

logrank P = 2e-7

S1	151	132	108	82	52	42
S2	59	43	31	23	16	10
S3	47	31	25	19	14	10
S4	20	10	4	3	3	2

d

NSCLC Radiotherapy (n=502)

logrank P = 1e-6

S1	147	105	54	30	15	2
S2	195	114	44	19	7	2
S3	111	42	19	9	3	3
S4	49	24	8	3	3	1
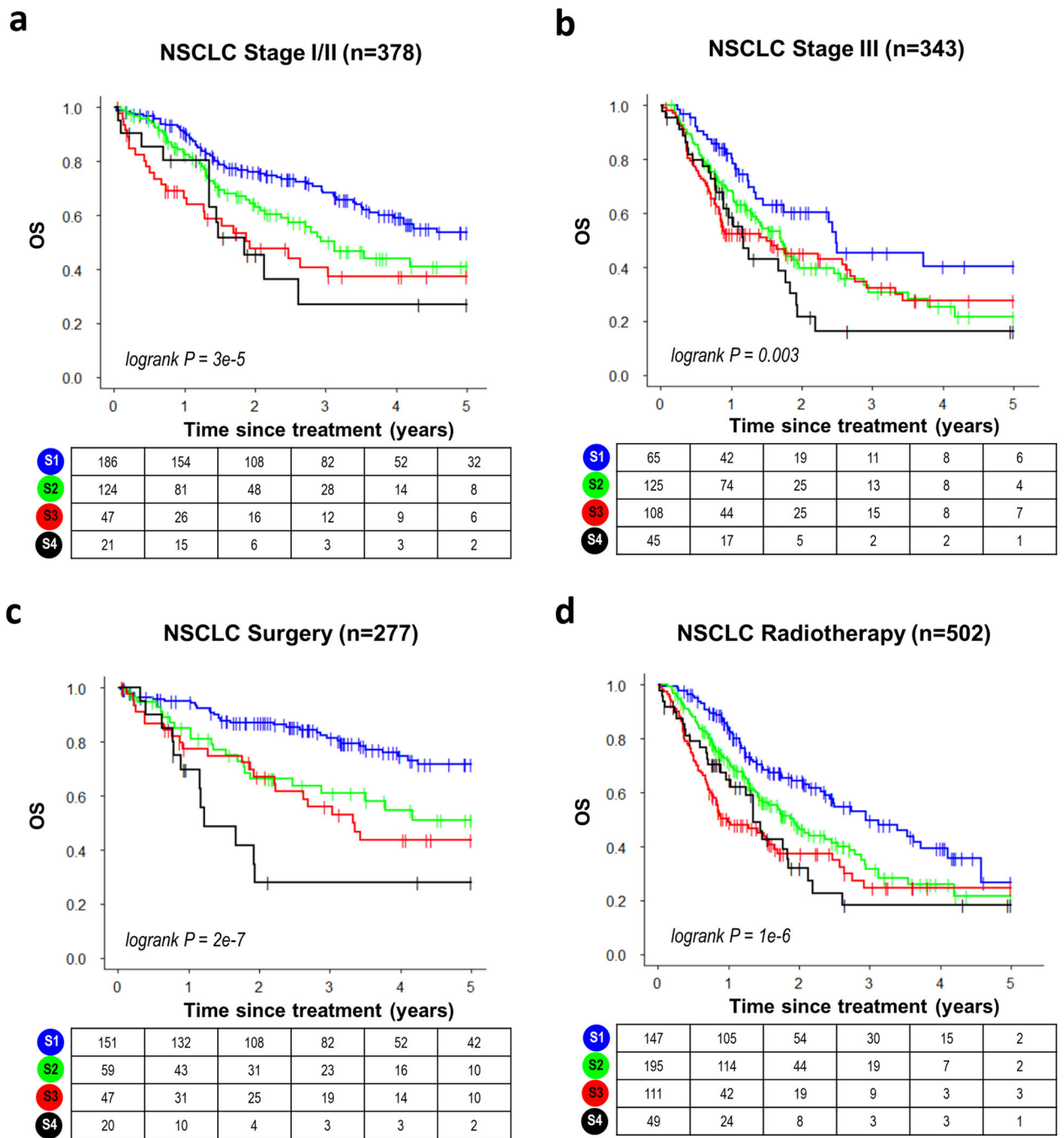
Figure 6.
Evaluation of prognostic value of the four imaging subtypes in lung cancer subgroups. Kaplan-Meier curves for a) stage I+II; b) Stage III; c) Patients treated with surgery; d) Patients treated with radiation.
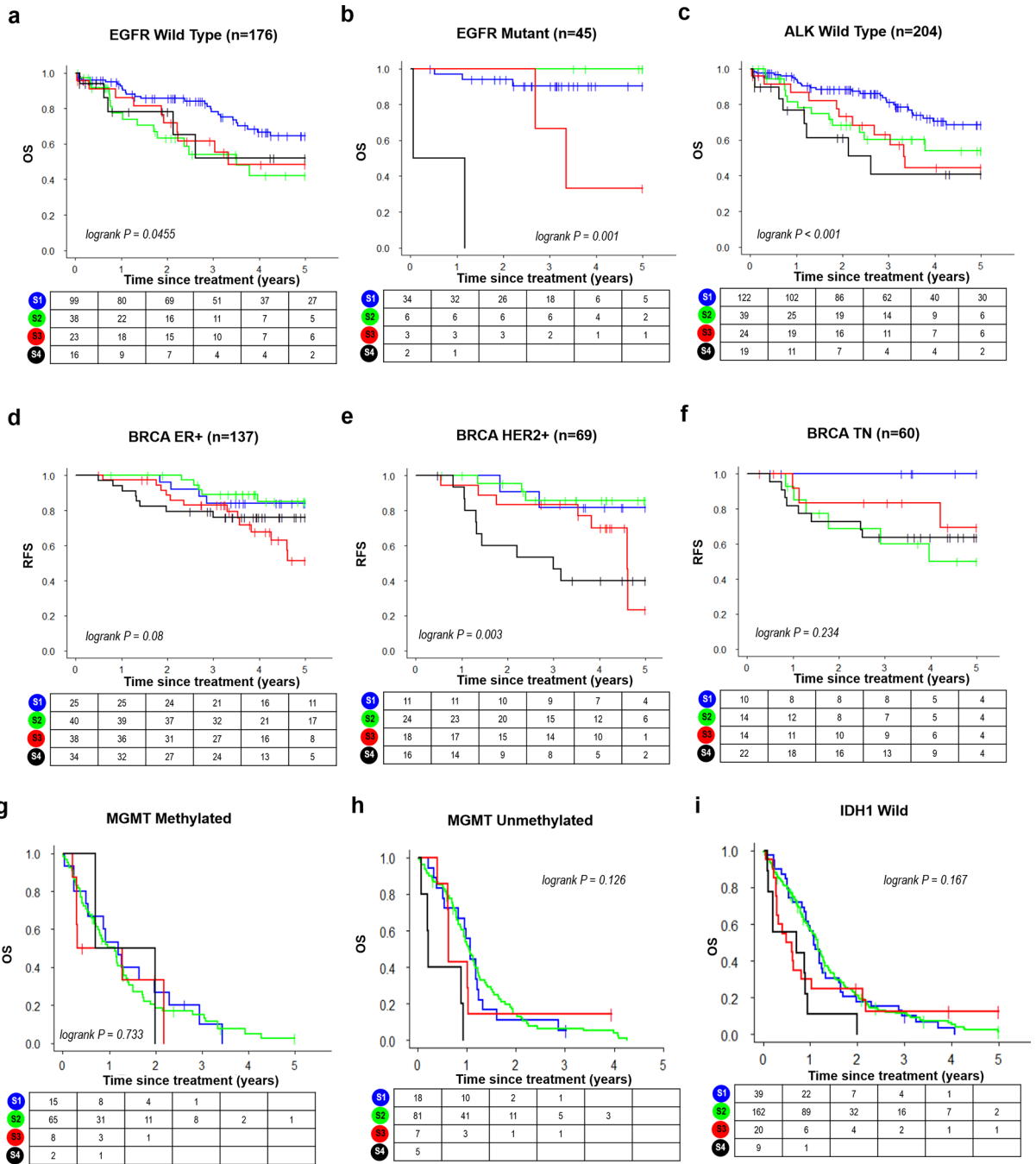
Figure 7.
Evaluation of prognostic value of the four imaging subtypes in subgroups within three cancer types. Kaplan-Meier curves for lung cancer subgroups: a) EGFR Wild Type; b) EGFR Mutant; c) ALK Wild Type; for breast cancer subgroups: d) ER+ group; e) HER2+ group; f) Triple Negative (TN) group; for GBM cancer subgroups: g) MGMT Methylated group; h) MGMT Unmethylated group; i) IDH1 Wild group.
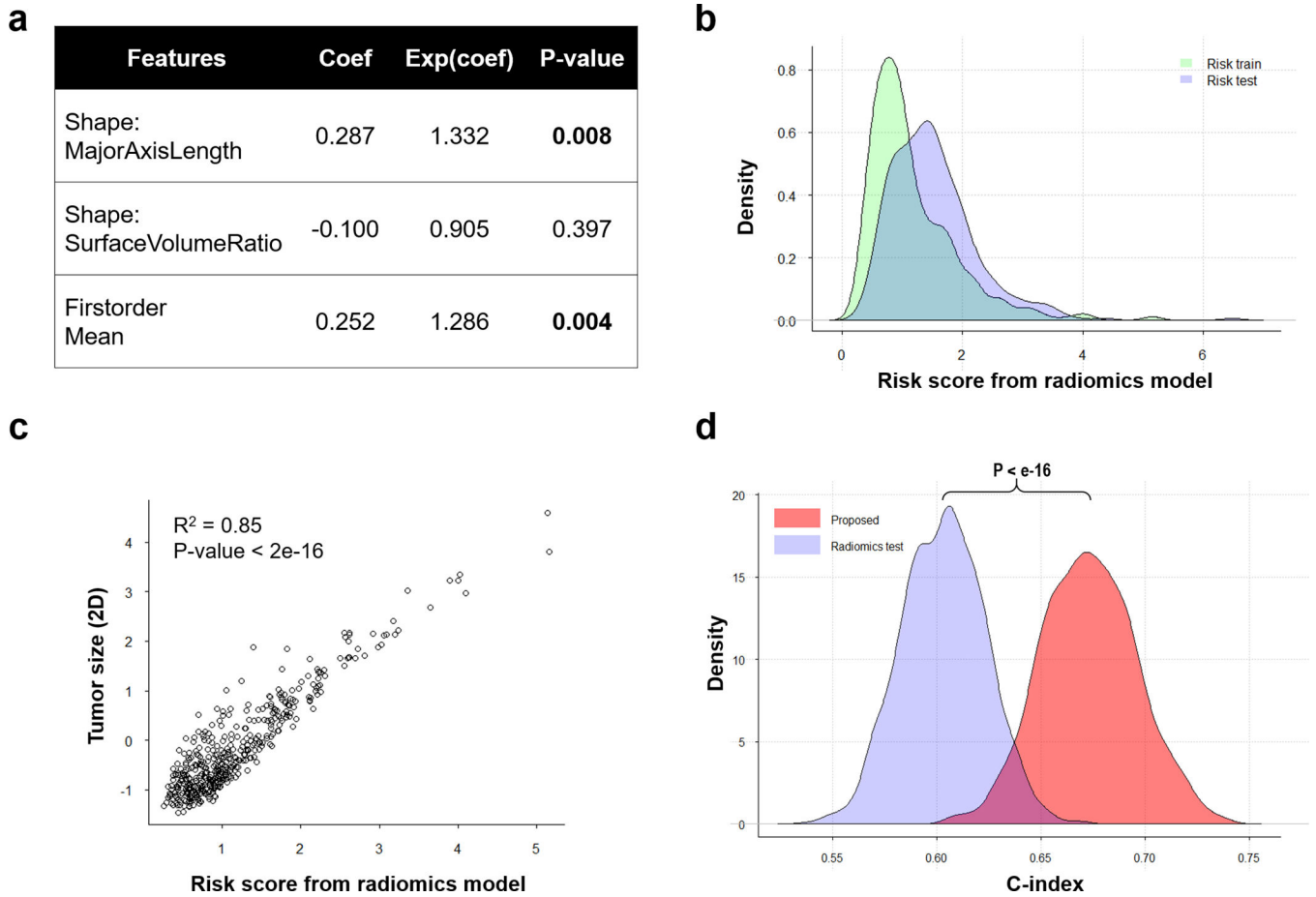
a

Features	Coef	Exp(coef)	P-value
Shape: MajorAxisLength	0.287	1.332	**0.008**
Shape: SurfaceVolumeRatio	-0.100	0.905	0.397
Firstorder Mean	0.252	1.286	**0.004**

b

c

d

Figure 8.
Comparison between the proposed imaging subtypes and conventional radiomics analysis for survival prediction in lung cancer cohorts. a) Details of the final radiomic model; b) Distribution of the radiomic risk score in training and validation cohorts; c) Scatterplot shows the correlation between radiomic risk score and tumor size measured in 2D; d) Distribution and comparison of c-index for the radiomic signature and the proposed imaging subtypes in the validation cohort.
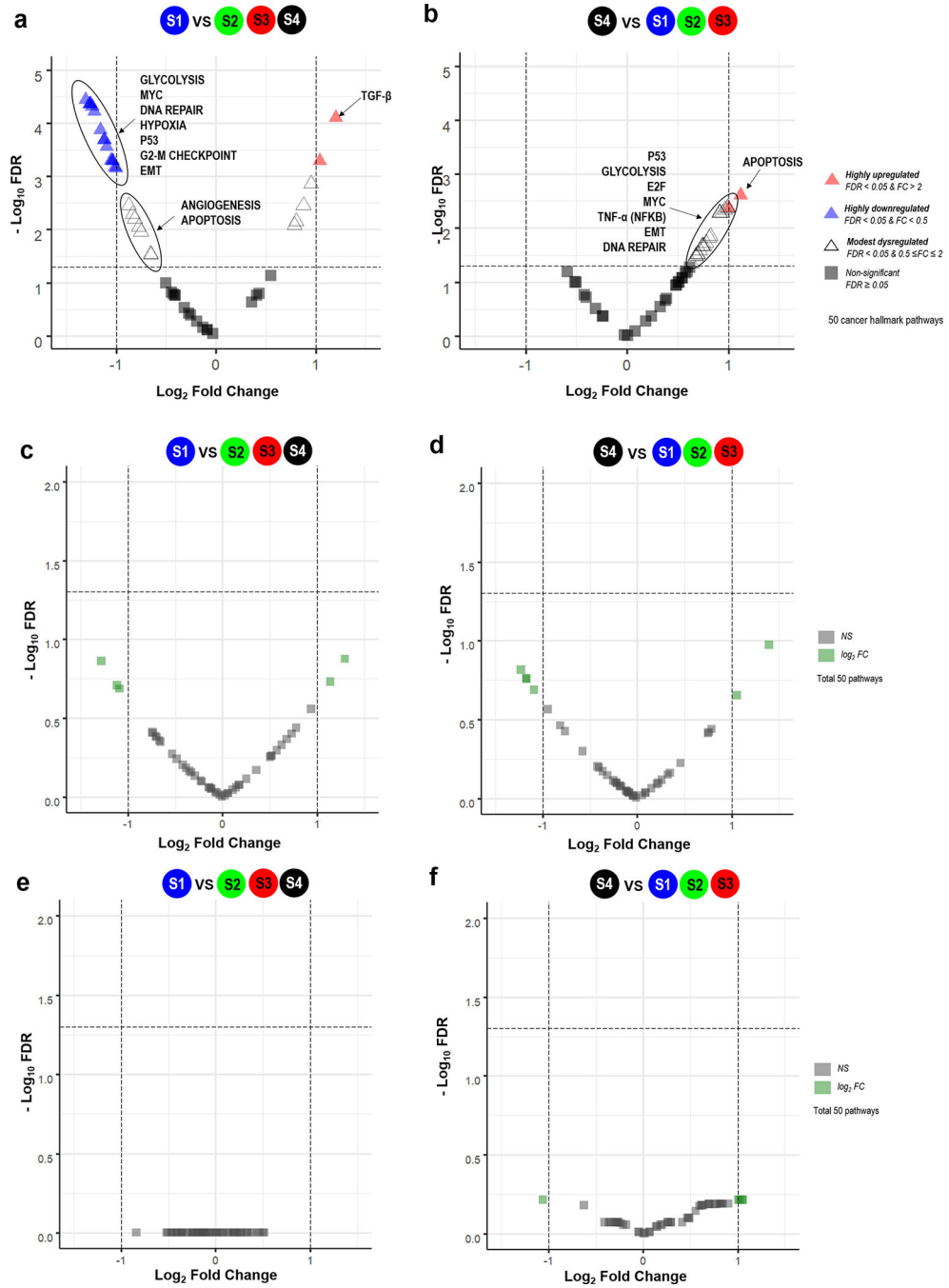
Figure 9.
Oncogenic processes associated with the imaging subtypes in three cancer types. Limma-modeled enrichment analysis by single-sample Gene Set Enrichment Analysis (ssGSEA) of 50 cancer hallmark pathways is applied. Volcano plot of enrichment scores in lung cancer: a) subtype 1 versus rest, and b) subtype 4 versus rest; in breast cancer: c) subtype 1 versus rest, and d) subtype 4 versus rest; in GBM: e) subtype 1 versus rest, and f) subtype 4 versus rest. The enrichment scores of 50 cancer hallmark pathways are plotted as log2 fold change versus the −log10 of the adjusted p-value. Thresholds are shown as dashed lines. Pathways

deemed as significantly different (false discovery rate [FDR] < 0.05) are highlighted with different color schemes.

| S1+S2 | 13 | 11 | 9 | 8 | 7 | 6 | 5 | 3 | 1 | | | | |
| S4 | 51 | 46 | 36 | 30 | 25 | 19 | 15 | 12 | 10 | 6 | 5 | 5 | 4 |
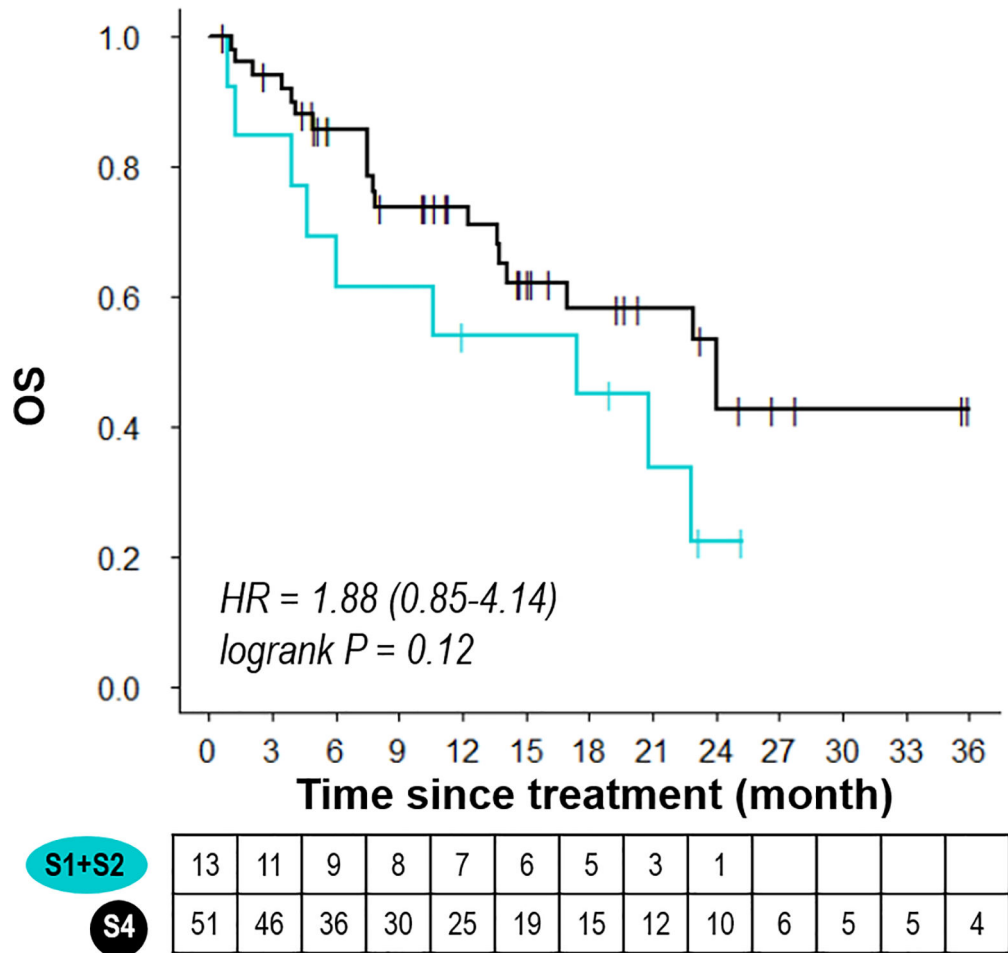
Figure 10.
Evaluation of imaging subtypes in the advanced lung cancer treated with immunotherapy. Kaplan-Meier curves of overall survival stratified by imaging subtype 1 and 2 versus 4.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References:

1. Lambin P et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48, 441–446 (2012). [PubMed: 22257792]

2. Gillies RJ, Kinahan PE & Hricak H Radiomics: images are more than pictures, they are data. Radiology 278, 563–577 (2015). [PubMed: 26579733]

3. Itakura H et al. Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities. Sci Transl Med 7, 303ra138 (2015).

4. Sun R et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. Lancet Oncol 19, 1180–1191 (2018). [PubMed: 30120041]

5. Jiang Y et al. Noninvasive imaging evaluation of tumor immune microenvironment to predict outcomes in gastric cancer. Ann. Oncol 31, 760–768 (2020). [PubMed: 32240794]

6. Vaidya P et al. CT derived radiomic score for predicting the added benefit of adjuvant chemotherapy following surgery in Stage I, II resectable Non-Small Cell Lung Cancer: a retrospective multi-cohort study for outcome prediction. Lancet Digit Health 2, e116–e128 (2020).

7. 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. Nature communications 11, 1–12 (2020).

8. Wu J et al. Magnetic resonance imaging and molecular features associated with tumor-infiltrating lymphocytes in breast cancer. Breast Cancer Research 20, 1–15 (2018). [PubMed: 29291743]

9. Berenguer R et al. Radiomics of CT Features May Be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. Radiology, 172361 (2018).

10. Mackin D et al. Measuring Computed Tomography Scanner Variability of Radiomics Features. Invest Radiol 50, 757–765 (2015). [PubMed: 26115366]

11. Traverso A, Wee L, Dekker A & Gillies R Repeatability and Reproducibility of Radiomic Features: A Systematic Review. Int. J. Radiat. Oncol. Biol. Phys 102, 1143–1158 (2018). [PubMed: 30170872]

12. Limkin E et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Annals of Oncology 28, 1191–1206 (2017). [PubMed: 28168275]

13. Lambin P et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 14, 749 (2017). [PubMed: 28975929]

14. Hoadley KA et al. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell 173, 291–304 e296 (2018). [PubMed: 29625048]

15. Pestana RC, Sen S, Hobbs BP & Hong DS Histology-agnostic drug development — considering issues beyond the tissue. Nature Reviews Clinical Oncology (2020).

16. O'Connor JPB et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 14, 169–186 (2017). [PubMed: 27725679]

17. Wu J, Mayer AT & Li R in Seminars in Cancer Biology (Elsevier, 2020).

18. Chalkidou A, O'Doherty MJ & Marsden PK False discovery rates in PET and CT studies with texture features: a systematic review. PLoS One 10, e0124165 (2015). [PubMed: 25938522]

19. Zhang YJ Geometric modeling and mesh generation from scanned images. (CRC Press, 2018).

20. Wu J et al. Intratumoral spatial heterogeneity by perfusion MR imaging predicts recurrence-free survival in locally advanced breast cancer treated with neoadjuvant chemotherapy. Radiology 288, 26–35 (2018). [PubMed: 29714680]

21. Braman NM et al. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. Breast Cancer Research 19, 1–14 (2017). [PubMed: 28052757]

22. Wu J et al. Robust intra-tumor partitioning to identify high-risk subregions in lung cancer: a pilot study. International Journal of Radiation Oncology* Biology* Physics 95, 1504–1512 (2016).

23. Yankeelov TE et al. Clinically relevant modeling of tumor growth and treatment response. Sci Transl Med 5, 187ps189 (2013).

24. Wu J et al. Tumor subregion evolution-based imaging features to assess early response and predict prognosis in oropharyngeal cancer. Journal of Nuclear Medicine 61, 327–336 (2020). [PubMed: 31420498]

25. Syed AK, Whisenant JG, Barnes SL, Sorace AG & Yankeelov TE Multiparametric analysis of longitudinal quantitative MRI data to identify distinct tumor habitats in preclinical models of breast cancer. Cancers 12, 1682 (2020).

26. Welch ML et al. Vulnerabilities of radiomic signature development: The need for safeguards. Radiotherapy and Oncology 130, 2–9 (2019). [PubMed: 30416044]

27. Cristescu R et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science 362 (2018).

28. Zhang YJ, Jing YM, Liang XH, Xu GL & Dong L Dynamic lung modeling and tumor tracking using deformable image registration and geometric smoothing. Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications Iii, 215–220 (2012).

29. Shukla-Dave A et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. J Magn Reson Imaging 49, e101–e121 (2019). [PubMed: 30451345]

30. Lawson DA, Kessenbrock K, Davis RT, Pervolarakis N & Werb Z Tumour heterogeneity and metastasis at single-cell resolution. Nature cell biology 20, 1349–1360 (2018). [PubMed: 30482943]

31. Lou B et al. An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction. The Lancet Digital Health 1, e136–e147 (2019). [PubMed: 31448366]

32. Coudray N et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med 24, 1559–1567 (2018). [PubMed: 30224757]

33. Jiang Y et al. Radiographic assessment of tumor stroma and treatment outcomes using deep learning: a retrospective multicohort study. Lancet Digit Health (2021).

34. Li A, Chen R, Farimani AB & Zhang YJ Reaction diffusion system prediction based on convolutional neural network. Scientific reports 10, 1–9 (2020). [PubMed: 31913322]

35. Li A, Farimani AB & Zhang YJ Deep learning of material transport in complex neurite networks. Scientific Reports 11, 1–13 (2021). [PubMed: 33414495]

36. Tajdari M et al. Image-based modelling for Adolescent Idiopathic Scoliosis: Mechanistic machine learning analysis and prediction. Computer methods in applied mechanics and engineering 374, 113590 (2021).

37. Kickingereder P et al. Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. The Lancet. Oncology 20, 728–740 (2019). [PubMed: 30952559]

38. Barajas RF et al. Regional variation in histopathologic features of tumor specimens from treatment-naive glioblastoma correlates with anatomic and physiologic MR Imaging. Neuro-Oncology 14, 942–954 (2012). [PubMed: 22711606]

39. Reynolds AP, Richards G, de la Iglesia B & Rayward-Smith VJ Clustering rules: a comparison of partitioning and hierarchical clustering algorithms. Journal of Mathematical Modelling and Algorithms 5, 475–504 (2006).

40. Kapp AV & Tibshirani R Are clusters found in one dataset present in another dataset? Biostatistics 8, 9–31 (2007). [PubMed: 16613834]

41. Thorsson V et al. The immune landscape of cancer. Immunity 48, 812–830. e814 (2018). [PubMed: 29628290]

42. Barbie DA et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462, 108–U122 (2009). [PubMed: 19847166]

43. Ritchie ME et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research 43, e47–e47 (2015). [PubMed: 25605792]

44. Ronneberger O, Fischer P & Brox T in International Conference on Medical image computing and computer-assisted intervention 234–241 (Springer, 2015).
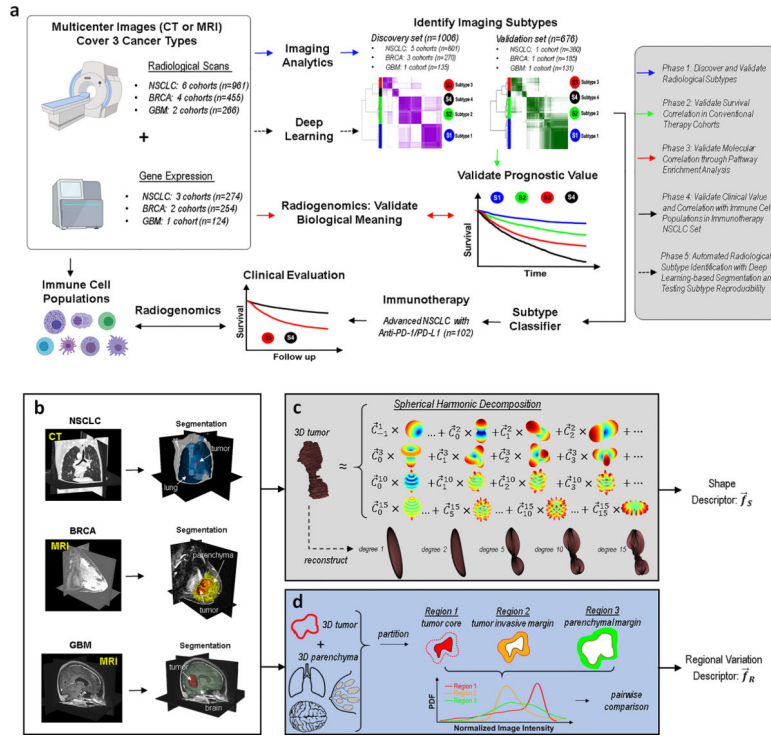
45. 10.5281/zenodo.4906510.

Figure 1.

Overview of the study design and quantitative imaging analysis. a) Study design, which contains five phases; b-d) Illustration of the proposed image feature extraction pipeline. First, the primary tumor was manually delineated and surrounding parenchymal tissues (i.e., lung, fibro-glandular, and brain) were automatically segmented, b). Then, two broad categories of image features were calculated, including c), systematic shape descriptors through spherical harmonic decomposition and d) spatial heterogeneity described by regional variations among tumor core, tumor invasive margin, and parenchymal region.
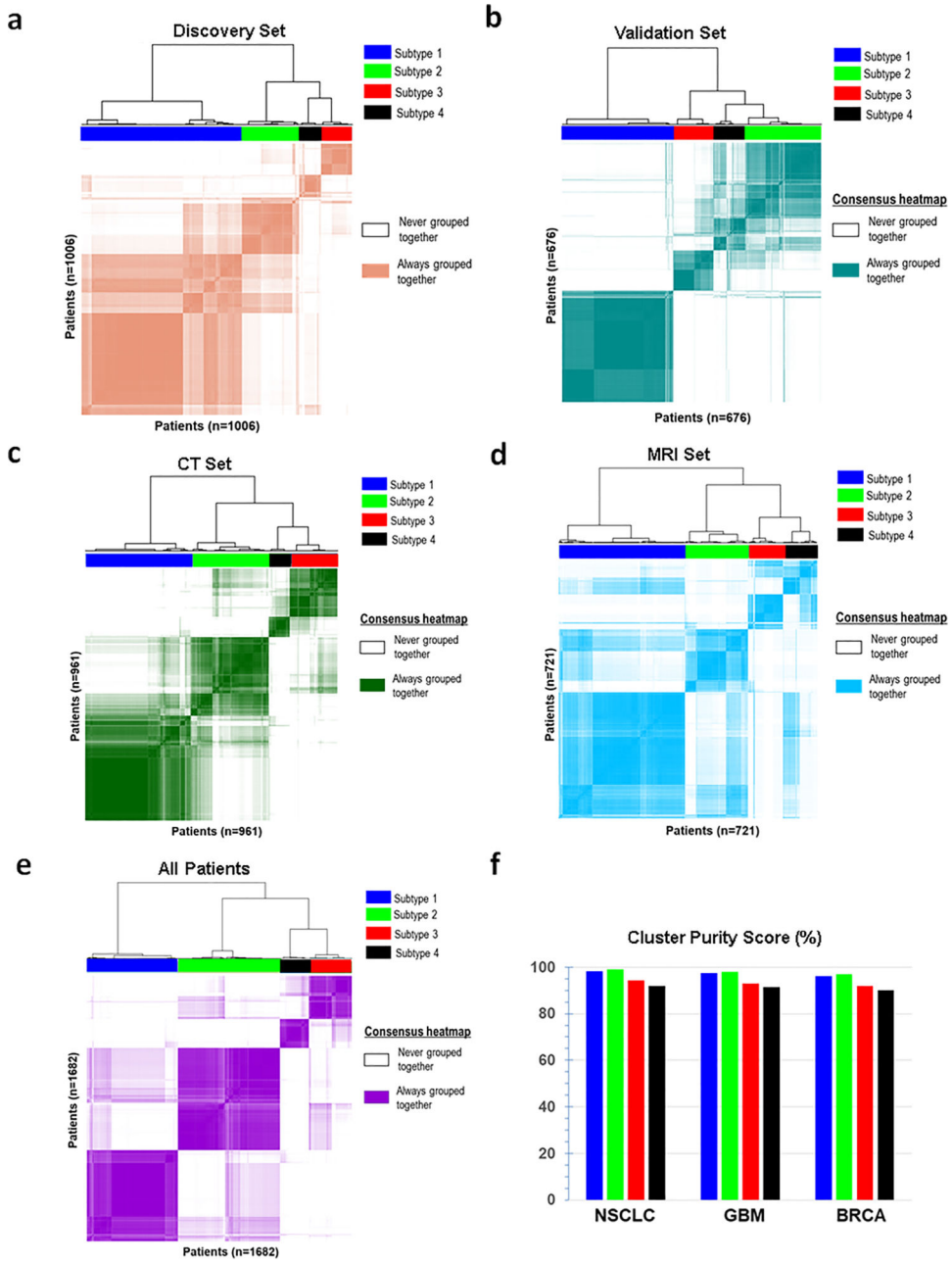
Figure 2.
Identification of unifying tumor subtypes based on unsupervised consensus clustering of the extracted image features across three cancer types and across two modalities (CT and MRI). The consensus matrix corresponding to the optimal cluster number (k=4) for a) discovery set, b) validation set, c) CT set, d) MRI set, and e) whole population. Patients are both rows and columns. The matrix is ordered by consensus-clustered groups, depicted as a dendrogram above the heat map. f) The cluster purity score of four tumor subtypes in three individual cancer types.
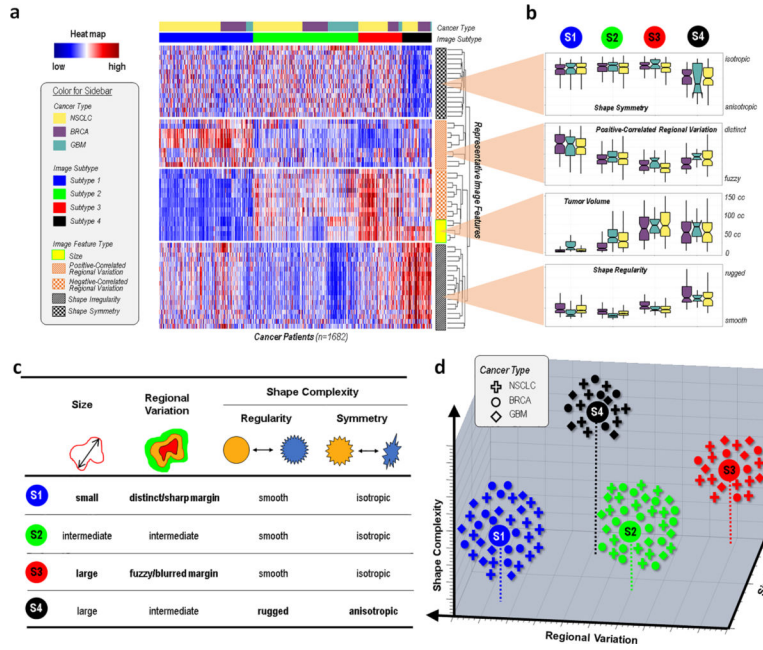
Figure 3.
Radiological characteristics of the unifying tumor subtypes. a) Heatmap of four subtypes with respect to the original imaging features; b) Boxplots of four representative groups of features including tumor volume, positive-correlated regional variation, shape symmetry, and shape irregularity, stratified by imaging subtypes as well as cancer types; c) Summary of key imaging characteristics of four subtypes; d) Schematic diagram for distribution of the imaging subtypes in a 3D space formed by tumor size, shape complexity, and regional variation.
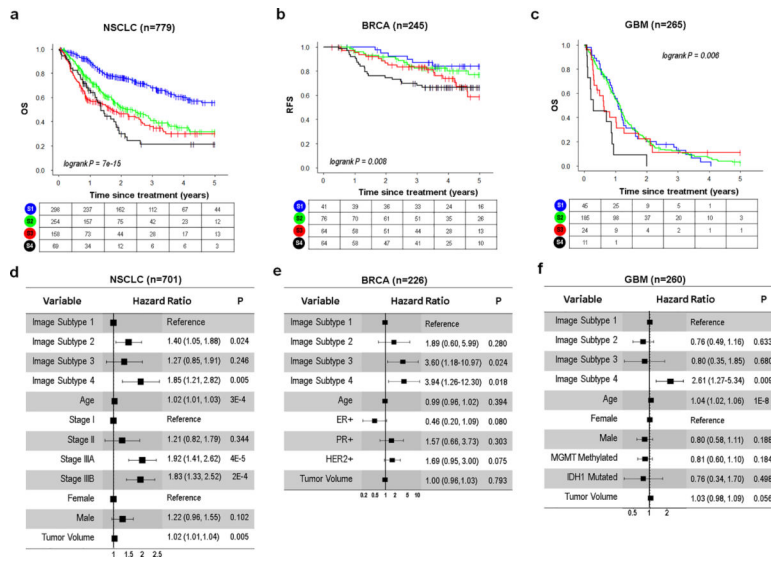
Figure 4.
Evaluation of prognostic value of the four imaging subtypes in three individual cancer types. Kaplan-Meier curves for a) overall survival in lung cancer, b) recurrence-free survival in breast cancer, and c) overall survival in glioblastoma multiforme. d-f) Forest plots show the hazard ratio and p values obtained from a multivariate Cox regression analysis including the proposed imaging subtypes and established clinicopathologic factors in different cancer types.
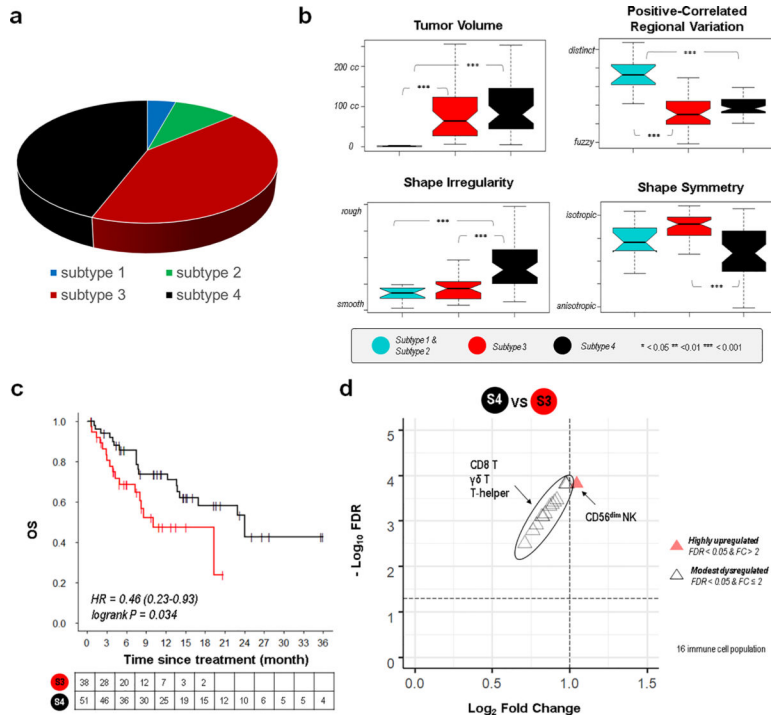
Figure 5.
Clinical evaluation of the imaging subtypes in advanced lung cancer treated with immunotherapy. a) Distribution of inferred imaging subtypes; b) Representative features stratified by imaging subtypes, including tumor volume, positive-correlated regional variation, shape symmetry, and shape irregularity; c) Kaplan-Meier curves of overall survival stratified by imaging subtype 3 versus 4; d) Comparison of tumor-infiltrating immune cell populations between subtypes 4 and 3.
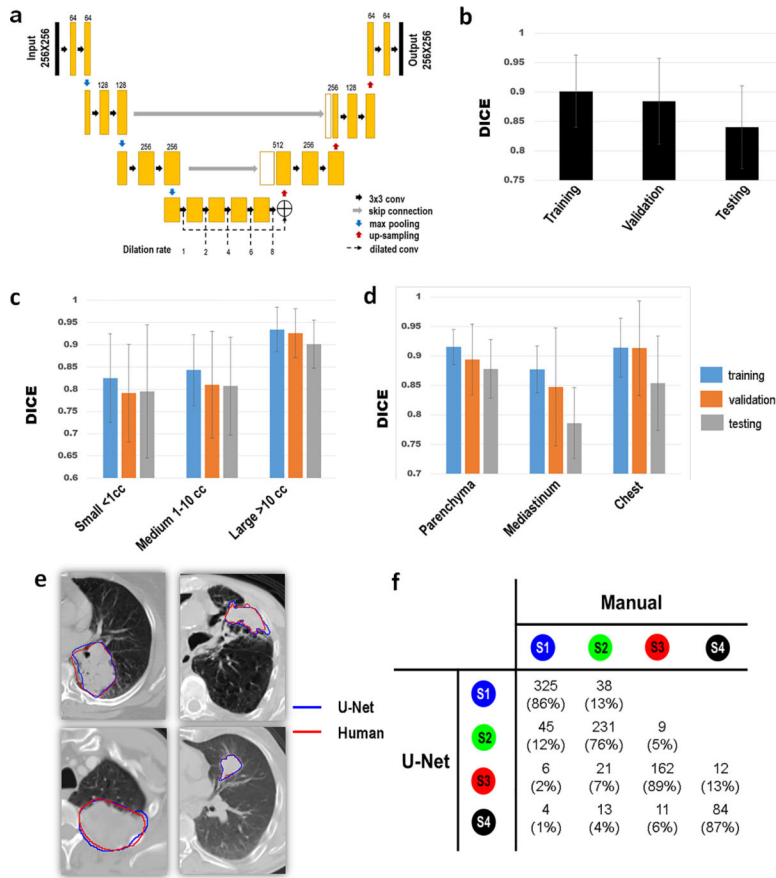
Figure 6.
Deep learning to automate 3D tumor segmentation. a) The detailed architecture of proposed U-Net with dilation at bottleneck layers; b) DICE coefficients of trained U-Net model in NSCLC in training, validation, and testing sets; DICE coefficients of U-Net stratified by c) tumor sizes, and d) anatomical locations of the tumor; e) Representative CT slices and tumor contours of four different testing patients; f) Confusion matrix of clustering results based on manual tumor contouring and automated segmentation for NSCLC patients.