



Quantification of Heterogeneity as a Biomarker in Tumor Imaging: A Systematic Review

Lejla Alic^{1,2*}, Wiro J. Niessen^{1,3}, Jifke F. Veenland¹

1 Biomedical Imaging Group Rotterdam, Department of Radiology and Medical Informatics, Erasmus Medical Center Rotterdam, Rotterdam, The Netherlands, **2** Department of Intelligent Imaging, Netherlands Organization for Applied Scientific Research (TNO), The Hague, The Netherlands, **3** Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

Abstract

Background: Many techniques are proposed for the quantification of tumor heterogeneity as an imaging biomarker for differentiation between tumor types, tumor grading, response monitoring and outcome prediction. However, in clinical practice these methods are barely used. This study evaluates the reported performance of the described methods and identifies barriers to their implementation in clinical practice.

Methodology: The Ovid, Embase, and Cochrane Central databases were searched up to 20 September 2013. Heterogeneity analysis methods were classified into four categories, i.e., non-spatial methods (NSM), spatial grey level methods (SGLM), fractal analysis (FA) methods, and filters and transforms (F&T). The performance of the different methods was compared.

Principal Findings: Of the 7351 potentially relevant publications, 209 were included. Of these studies, 58% reported the use of NSM, 49% SGLM, 10% FA, and 28% F&T. Differentiation between tumor types, tumor grading and/or outcome prediction was the goal in 87% of the studies. Overall, the reported area under the curve (AUC) ranged from 0.5 to 1 (median 0.87). No relation was found between the performance and the quantification methods used, or between the performance and the imaging modality. A negative correlation was found between the tumor-feature ratio and the AUC, which is presumably caused by overfitting in small datasets. Cross-validation was reported in 63% of the classification studies. Retrospective analyses were conducted in 57% of the studies without a clear description.

Conclusions: In a research setting, heterogeneity quantification methods can differentiate between tumor types, grade tumors, and predict outcome and monitor treatment effects. To translate these methods to clinical practice, more prospective studies are required that use external datasets for validation: these datasets should be made available to the community to facilitate the development of new and improved methods.

Citation: Alic L, Niessen WJ, Veenland JF (2014) Quantification of Heterogeneity as a Biomarker in Tumor Imaging: A Systematic Review. PLoS ONE 9(10): e110300. doi:10.1371/journal.pone.0110300

Editor: Christos Hatzis, Yale University, United States of America

Received: February 26, 2014; **Accepted:** September 15, 2014; **Published:** October 20, 2014

Copyright: © 2014 Alic et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by The Netherlands Organization for Scientific Research (NWO), grant number 017.002.019. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: LejlaResearch@gmail.com

Introduction

Tumors are often inhomogeneous. Regional variations in cell death, metabolic activity, proliferation and vascular structure are observed. There is increasing evidence that solid tumors may consist of subpopulations of cells with different genotypes and phenotypes [1]. These distinct populations of cancer cells can interact in a competitive way [2] and may differ in sensitivity to treatments [3,4]. This heterogeneity can be detected using diagnostic imaging techniques at a genetic, molecular or cellular level [4,5], or at a cell population level. The advantage of diagnostic imaging techniques is their non-invasive nature and the fact that the whole tumor is taken into account, whereas cellular diagnostic techniques are invasive and limited to a discrete set of tumor samples. Various imaging techniques are available to visualize the heterogeneity in tissue characteristics, such as necrosis, metabolic activity, cell density and vascularity. Observed heterogeneity in an image is a reflection of the phenotypic

variation of the tumor and is reported to be associated with underlying gene-expression patterns [6].

Image heterogeneity can be quantified using a variety of texture analysis methods. As such, image heterogeneity is potential biomarker for tumor characterization, for response prediction and monitoring. Parameters in hot spots, as quantified with dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), are more relevant for monitoring tumor response than parameters averaged over the whole tumor [7–9]. When a region of the tumor is not well vascularized or is hypoxic, chemotherapy and radiotherapy are more likely to fail. The existence of poorly vascularized or hypoxic areas within a tumor is an important component of tumor radiation resistance and correlates with treatment failure [10]. In radiotherapy, the heterogeneity can be used to guide treatment [11,12]: an ongoing trial is currently escalating the dose to the part of the tumor with high standardized uptake values [13]. Also for computed tomography (CT), image heterogeneity has prognostic value [6].

Several methods are available to quantify tumor heterogeneity from imaging data. Many studies have used histogram-derived features such as percentile values, standard deviation (SD) and enhancing fraction. However, these features do not take into account the spatial distribution of the intensity values. In contrast, texture methods take spatial information into account by quantifying the spatial variations in the images. Ideally, these methods are independent of the absolute signal intensities in the image. They provide additional and independent information (such as the average signal intensity) compared to histogram-derived measures. These methods result in features which can be considered to be imaging biomarkers providing information on the underlying tumor heterogeneity. Some of these features are related to image properties that are visually perceived by the radiologist, whereas others are more abstract [14].

By means of a systematic review, the aim of this study is to investigate the performance of different heterogeneity imaging biomarkers extracted from diagnostic tumor images for differentiation between tumor types, tumor grading, outcome prediction and treatment monitoring.

The following research questions were formulated:

- Which analysis methods are used to quantify heterogeneity or texture in tumor imaging, with the aim to differentiate between tumor types, tumor grading, outcome prediction and treatment monitoring?
- What are the reported performances of the different analysis methods? Is there a relation between performance and analysis method?

What is the potential clinical impact of the methods? Can the performance results be generalized? Is the performance evaluated in addition to established imaging biomarkers?

Methods

Data Sources and Search method

This review was performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses) guidelines [15], with details summarized in Checklist S1. In January 2013 the study protocol was registered with the International Prospective Register of Systematic Reviews (Identification number: CRD42013003634) [16]. A systematic search was conducted in the databases of Medline, Embase, and Cochrane Central. The search was performed with the aid of an experienced librarian on September 20th 2013.

The following topics were used for the searches:

1. Neoplasms
2. Heterogeneity, texture
3. MRI, MRS, CT, PET, SPECT, ultrasonography
4. Differentiation between tumor types, tumor grading, classification, staging, treatment response, survival, and treatment outcome

Full details of the Embase search is included in Text S1. The results from all three searches were combined and verified to ensure exclusion of publications containing the same title, written by the same authors, and published in the same journal. The remaining publications were considered for study selection.

Study Selection

Two authors (L.A. and J.F.V.) independently reviewed the titles and abstracts. The selected publications then underwent full-text

screening. During the title and abstract review, any discrepancies about study inclusion were resolved by full-text screening. Any discrepancies during the following stages were resolved by discussion. The bibliographies of seminal review papers [17–19] were reviewed to identify additional relevant articles.

Inclusion and exclusion criteria

We included only publications related to diagnostic imaging which reported quantification of tumor heterogeneity or tumor texture with the goal to differentiate between tumor types, tumor grading, outcome prediction and tumor response monitoring. No restrictions were made based on location, type, stage or grade of malignancy. Prior to review, a decision was made to exclude any study with too few participants, i.e., for patient studies ($n < 10$) and for animal studies ($n < 5$). Therefore, all case studies, and studies with no information on the number of subjects, were excluded. In addition, all the following types of studies were excluded:

- publications based on non-tumor images;
- publications not based on quantitative assessment of heterogeneity or texture in images
- publications without one of the following goals: differentiation between tumor types, tumor grading, or outcome prediction or treatment monitoring;
- publications not based on in vivo studies (histology, phantom, ex vivo, synthetic data);
- publications describing non-original research (editorial, letter to the editor, review, meta-analysis, opinion publications).

Data extraction

A data extraction form was designed. All selected publications were independently reviewed and data extraction was cross-checked. Disagreements between the reviewers were resolved by consensus. The following data were extracted from the full papers: year of publication, human or animal study, type of study (retrospective or prospective), number of subjects, number of tumors, location of tumor, imaging modality, tracer/contrast agent, goal of heterogeneity/texture analysis, and type of heterogeneity/texture quantification method used. For studies reporting on the same analysis method based on the identical dataset, only the latest publication was included. For publications reporting classification experiments, the following data were extracted: number of candidate heterogeneity features, dimensionally reduction technique used, number of selected features used in the best classification experiment, the results of the best classification experiment, i.e., accuracy, sensitivity, specificity, area under the receiver operator curve (AUC), type of cross-validation used, and use of an external validation set. For publications using statistical hypothesis testing the following data were extracted: the number of candidate features, and the number of features that showed a significant difference between outcome categories (before and after Holm-Bonferroni correction) [20]. All publications were divided into two categories:

- Publications reporting cross-sectional measurements with the aim to differentiate between tumor types, tumor grading, and treatment outcome prediction.
- Publications reporting longitudinal measurements for tumor treatment monitoring.

Data synthesis and analysis

The imaging modalities were summarized into four categories: i) magnetic resonance imaging (MRI), ii) computed tomography (CT), iii) positron emission tomography (PET), single photon emission computed tomography (SPECT), and iv) ultrasonography (US). No further subdivision was made regarding the type of imaging protocol or use of contrast agent.

Image analysis methods to estimate tumor heterogeneity were divided into four categories: non-spatial methods, local spatial distribution methods, fractal analysis, and a category consisting of filters and transforms.

Non-spatial methods (NSM). These methods characterize tumor heterogeneity by non-spatial descriptors, such as descriptors of the gray-level frequency distributions: standard deviation, skewness, maximum, minimum, range, peak height, peak position, and percentile values.

Spatial gray-level methods (SGLM). Methods included in the second category extract the local spatial image intensity distribution. This category includes grey-tone spatial-dependence matrix (GTSDM) [21], neighborhood gray-tone difference matrix (NGTDM) [22], run-length matrix (RLM), and Local Binary Pattern (LBP) [23]. The GTSDM, originally proposed by Haralick et al. [21], is often referred to as co-occurrence or the second-order histogram. When divided by the total number of neighboring pixels in the image, this matrix becomes the estimate of the joint probability of two pixels at a distance along a given direction having a particular gray value. The NGTDM, originally proposed by Amadasm and King [22], is based on spatial changes in gray values by inspecting the difference between gray levels of a specific pixel and the average gray level of their surrounding neighbors. The RLM, originally proposed by Galloway [24], is subsidiary to the observation that a coarse texture would have relatively longer gray level runs compared to a fine texture. This matrix provides information about runs of pixels with the same gray level values in a given direction. LBP, originally proposed by Ojala et al. [25] and later modified to a rotation and scale invariant approach [23], represents local texture. In its simplest form it labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number.

Fractal analysis (FA). The third category consists of FA methods that overcome the scale problem by providing a statistical measure reflecting pattern changes as a function scale. The two basic parameters in FA are fractal dimension (FD) and lacunarity [26]. An often used method to estimate FD is box counting [26]. This procedure systematically overlays an image with a series of grids with increasing/decreasing size. For each step, this procedure captures the predefined relevant features [27]. Another frequently used technique in FA is the blanket method [26], which is often used in its extended form, as described by Peleg et al. [28]. This method estimates the surface area by measuring the volume between an upper and lower blanket.

Filters and Transforms (F&T). The fourth category consists of a collection of image processing algorithms that extract texture features. Examples are methods that use techniques defined in the spatial domain such as filters (Gabor filters or Law's filters) or transformations to other domains (Fourier transform, Wavelet transform, S-transform, discrete cosine transform). Since the various methods have only been used in a limited number of publications included in the present review, these methods were grouped together.

Publications reporting classification experiments. Publications were considered classification studies if they reported a classification result such as accuracy, sensitivity, specificity or AUC values. Only publications in which

the results of the classification experiments were solely based on texture parameters were further analyzed. These studies often utilize a high number of candidate features to describe a tumor. When the number of extracted features is too large to perform a statistically meaningful classification [29], the extracted features can be redundant in the information they retain. Because an increase of dimensionality in the feature space results in an increase of its volume, the feature space is sparsely filled. The use of an extensive number of features for classification purposes can result in over-fitting, which reduces the possibility of generalization; this paradox is generally referred to as the 'curse of dimensionality' [30].

To keep the system manageable, dimensionality reduction techniques were commonly applied to select a subset of features that were relevant for the classification problem. The ratio between the number of tumors classified and the dimensionality of the feature space (e.g., the number of selected features) should be chosen in a meaningful way. In pattern recognition applications, the rule of thumb is to use 5–10 datasets per feature per category [31]. Therefore, we evaluated the number of candidate features, the number of selected features, and the ratio between the number of tumors included in the study and the number of selected classification features. A one-way ANOVA was used to test for differences in classification results between the modalities and analysis methods.

Publications reporting on significance testing. A commonly used approach to test the validity of the selected features is significance testing. For heterogeneity analysis, many publications compute a large number of features. As multiple comparisons generally require a stronger level of evidence to be considered significant, the Holm-Bonferroni correction [20] can be applied. This correction allows for the significance levels for single and multiple comparisons to be directly comparable. In these publications, we evaluated whether a Holm-Bonferroni correction was applied and, if this was not the case, computed the number of significant features after correction using the available data. A one-way ANOVA was used to test for differences in the number of significant features, before and after Holm-Bonferroni correction, between the modalities and the analysis methods used.

Results

Figure 1 presents details on the literature search. In summary, of the 7351 potentially relevant articles, 480 (6.5%) were considered for inclusion after abstract review. After these latter papers had undergone full-text screening, an additional 249 publications were excluded. The remaining 231 original publications entered the data extraction phase. In this phase an additional 22 papers [32–53] were excluded as they reported results of a similar analysis method on the same dataset as that used in another paper; for these publications, the most recent one was included in the analysis. Finally, data from 209 studies [7,14,54–228] were extracted for further analysis.

General characteristics

Table 1 presents the characteristics of the included publications (after removing duplicate publications). A publication may include more than one imaging modality, analysis method, or goal. Two studies (1%) reported on two imaging modalities, and 66 studies (32%) reported on two or more analysis methods.

Since 2008, the number of imaging studies quantifying tumor heterogeneity has been steadily increasing, i.e. from 8 papers in 2006–2007 to 66 publications in 2012–2013 (figure 2-A). Prior to 2006, heterogeneity was mainly studied based on US data

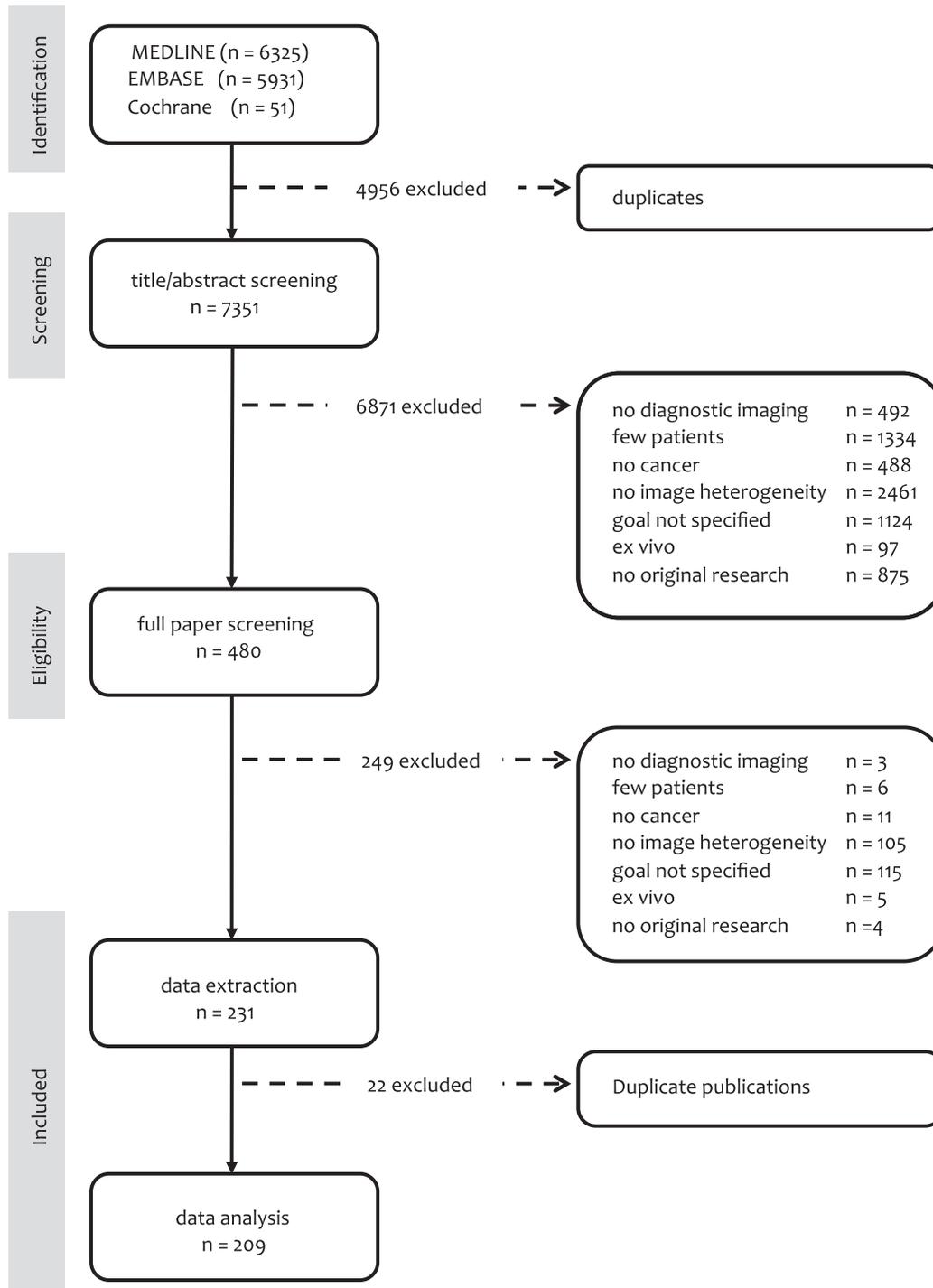


Figure 1. Results of the literature search. PRISMA flow diagram for study collection [15], showing the number of studies identified, screened, eligible, and included in the systematic review. This study is registered with the PROSPERO registry for systematic reviews (Identification number: CRD42013003634) [16]. doi:10.1371/journal.pone.0110300.g001

(Figure 2-B). Since 2007, most studies quantifying tumor heterogeneity are based on MRI. Generally, the non-spatial method (NSM) and the spatial gray-level method (SGLM) are the most frequently used to analyze tumor heterogeneity (Figure 2-C). Although the number of publications using these methods has increased since 2007, their contribution to heterogeneity literature is relatively stable. The number of studies reporting tumor

response monitoring has varied over the years, ranging from 0–20% (Figure 2-D).

Breast tumors were studied in 33% (n = 69) of the publications. Figure 3 shows the distribution of studies per tumor location. Figure 3-A shows the use of imaging modalities for quantification of tumor heterogeneity per primary tumor location. MRI is used primarily for brain and breast tumors, CT for lung and

Table 1. Characteristics of the included publications (n = 209).

Characteristic		n	%
Imaging method	MRI	75	36%
	CT	40	19%
	PET	14	7%
	US	81	39%
Analysis method	NSM	121	58%
	SGLM	103	49%
	FA	21	10%
	F&T	58	28%
Study goal	Diagnosis/grading/outcome pred.	182	87%
	Response monitoring	27	13%
Study type	Retrospective	118	56%
	Retrospective (with inclusion criteria)	63	30%
	Prospective	28	13%
Type of subjects	Human	197	94%
	Animal	12	6%
Type of experiment	Classification	139	67%
	Significance testing	64	30%
	Neither	6	3%

Imaging modalities: magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), ultrasonography (US).
 Analysis methods: non-spatial methods (NSM), spatial grey level methods (SGLM), fractal analysis (FA) methods, and filters and transforms (F&T).
 doi:10.1371/journal.pone.0110300.t001

gastrointestinal tumors, PET for gastrointestinal, lung tumors and sarcoma, and US for breast tumors. Heterogeneity analysis of brain tumors was performed almost exclusively with MRI, while for breast tumors both MRI and US were used.

Figure 3-B presents the analysis methods used per primary tumor location. For almost all locations, all methods were used. For prostate, breast, and head and neck analysis, the SGLM was the most frequently used. For all other locations, the NSM was the favored modality. Heterogeneity analyses for longitudinal studies were mainly performed for gastrointestinal and breast tumors (Figure 3-C).

Figure S1 summarizes the publications included in the present review (n = 209) in a matrix form. The publications are divided into different imaging modalities and analysis methods, and are available for download for each cell separately. Each cell in the matrix links to the supplementary EndNote file containing the records for these publications.

Figure 4-A shows the relation between imaging modality and analysis methods for cross-sectional studies. In general, 74% of these studies used either MRI or US. The SGLM (37%) and NSM (36%) are most frequently used to grade and diagnose tumors. Figure 4-B shows the relation between imaging modality and analysis method for the longitudinal studies (n = 27). MRI was

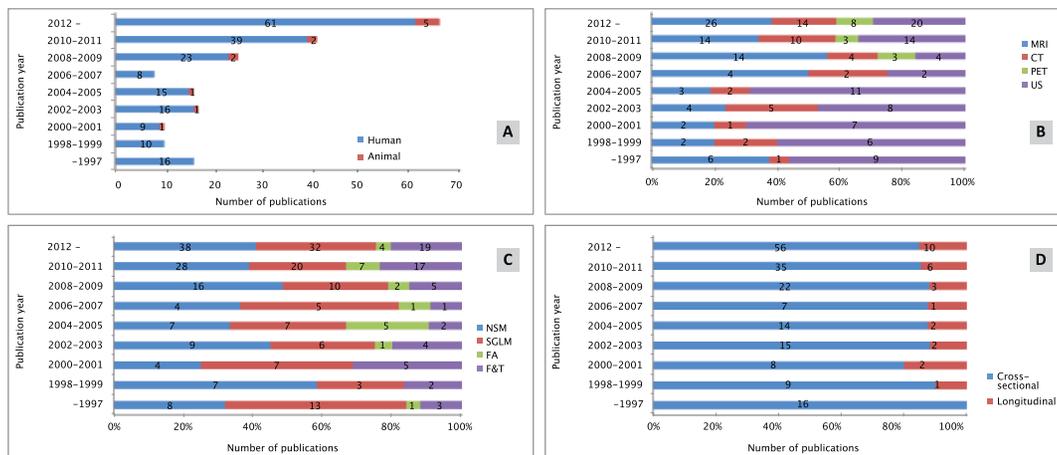


Figure 2. Number of publications reporting on tumor heterogeneity analysis for all publications bi-annually. Total number of publications (A), publications per imaging modality (B), publications per analysis method (C), and publications per goal (D).
 doi:10.1371/journal.pone.0110300.g002

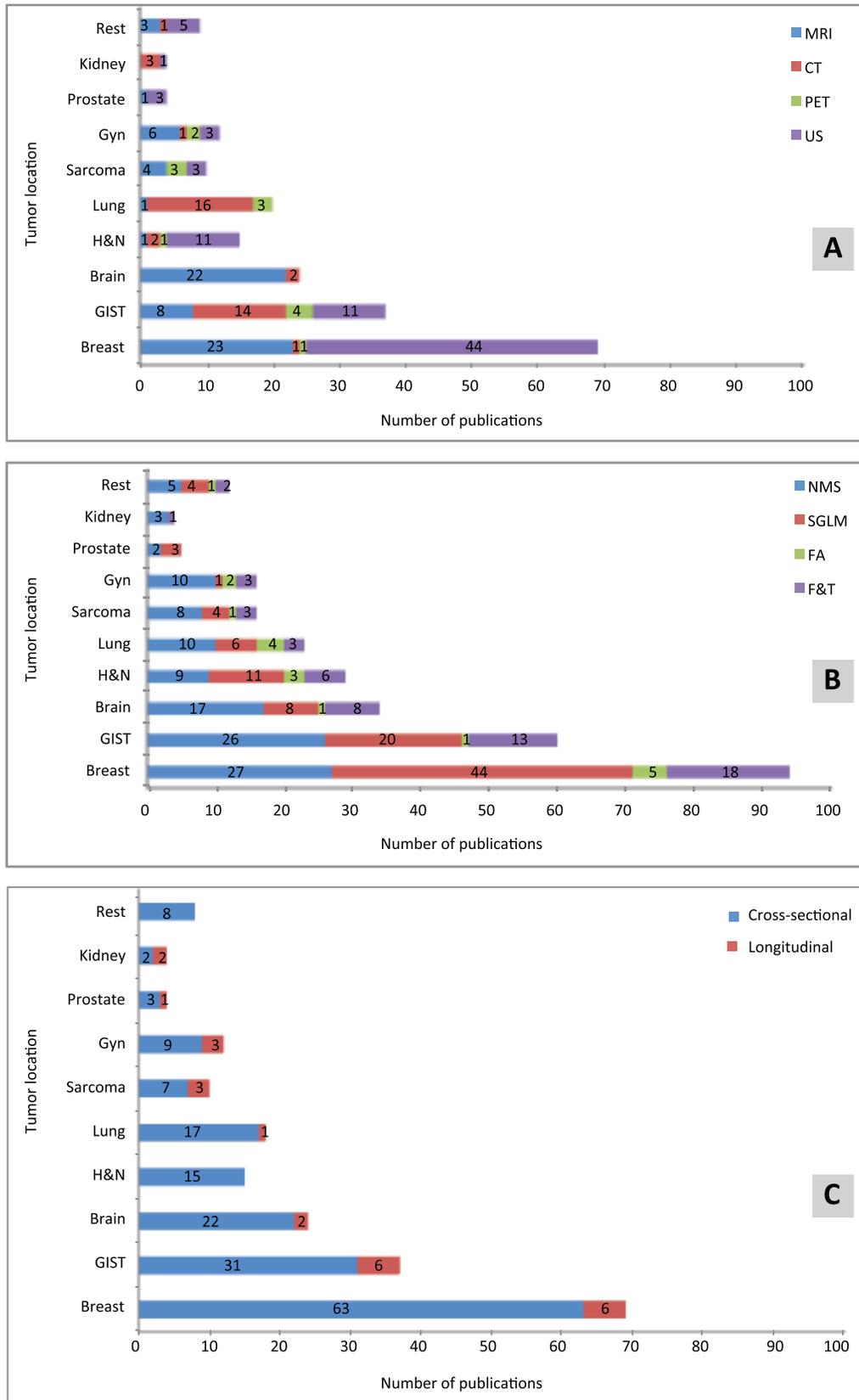


Figure 3. Publications reporting on quantification of tumor heterogeneity in cancer sites summarized for imaging modality (A), analysis method (B), and study aim (C). Publications can report on more than one analysis method. The acronyms used: Gyn – gynecological, H&N - head and neck, GIST – gastrointestinal.
doi:10.1371/journal.pone.0110300.g003

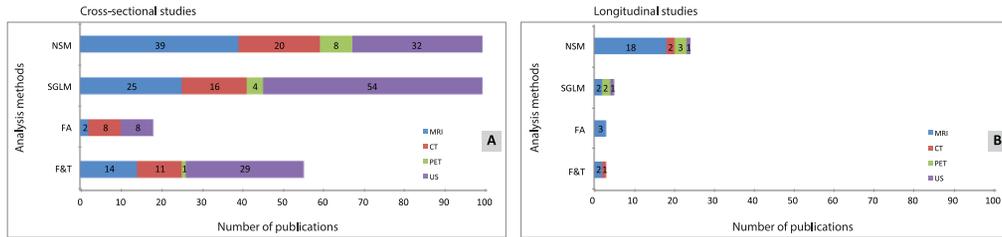


Figure 4. All included publications reporting cross-sectional (A) and longitudinal (B) studies. Several publications report more than one analysis method.
doi:10.1371/journal.pone.0110300.g004

used in 70% of these studies and PET in 11%. In 7% of the studies, US-based heterogeneity quantification was used for tumor response monitoring. NSM is the most frequently used (69%) analysis method in longitudinal studies.

A relatively small number of all studies (13%) utilized a prospective study design. Figure 5-A shows the relation between imaging modality and analysis method used for cross-sectional studies (n = 12). US is the most frequently used modality, whereas NSM is the most frequently used analysis method. Figure 5-B shows the relation between imaging modality and analysis method for publications reporting longitudinal studies (n = 16). Again, most data were analyzed with NSM. In contrast to MRI, CT, US and PET are rarely used for heterogeneity quantification in prospective longitudinal studies.

Publications reporting classification experiments

Of all included studies, 67% (n = 139) reported classification experiments and 30% reported significance testing. The remaining 3% either did not report quantitative results or the experiments were not completely described. Also, 23 studies only reported results of classification experiments where the texture features were combined with non-texture features. For these latter publications, it was not possible to extract the performance of the texture features separately and, therefore, these results were excluded from further analysis. Additionally, 10 studies were excluded because the number of generated or selected features was lacking. Of the papers reporting classification experiments (n = 106), 45% used US, 37% used MRI, 13% used CT, and 5% used PET. In 42% of the classification papers, features originating from different texture analysis methods were combined. Some studies reporting classification experiments (n = 39) performed no feature reduction, and the median number of candidate features used in these studies (6) was significantly lower than that of candidate features in the studies using feature reduction techniques (38). The remaining 67 studies reporting classification experiments used one of the

methods commonly applied in statistics, pattern recognition, or machine learning. These methods were summarized into three categories: filters, wrappers and embedded methods [229].

Figure 6 shows the relation between the number of candidate features and the number of selected features used in classification experiments for different imaging modalities (Figure 6-A) and different analysis methods (Figure 6-B). For the papers presented on the dotted line, no feature selection was performed. The number of candidate features ranged from 1–5280 (median 22) while the number of selected features ranged from 1–476 (median 3). The distribution of the numbers of selected features can be assessed as boxplots for imaging modality (Figure 6-C) and for analysis methods (Figure 6-D).

About 63% of the publications describing a classification experiment, reported cross-validation or training test sets as a technique to limit the effect of over-fitting on the available data. Figure 6-B shows that the combination of features from different methods generally leads to a higher number of candidate features. In general, in publications reporting the use of more than one analysis method more extensive feature reduction is applied compared to publications reporting on the use of the separate analysis methods.

In the classification experiments, one or more of the following performance measures were reported: sensitivity, specificity, accuracy, or AUC. Figure 7-A shows the AUC per imaging modality and Figure 7-B the AUC per analysis method. The differences in performance (as measured by AUC) are shown in Figure 7-C per imaging modality and in Figure 7-D per analysis method.

The supplementary material provides the figures for accuracy (Figure S2), sensitivity (Figure S3) and specificity (Figure S4) per imaging modality and per analysis method. In these figures, the reported performance is depicted as a function of the tumor-feature ratio (ratio between the number of tumors included and the number of selected features). In general, the tumor-feature

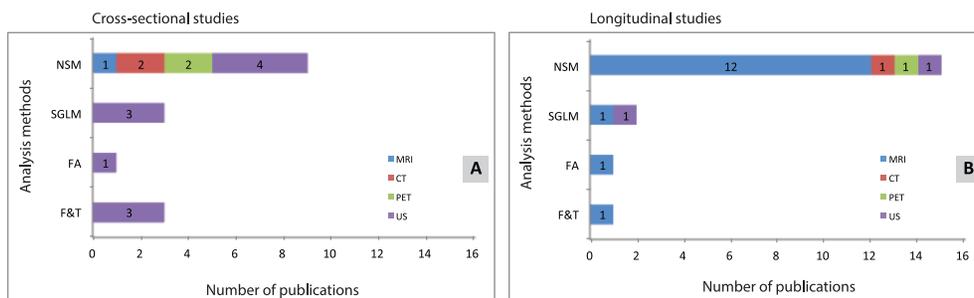


Figure 5. Publications reporting a prospective study design cross-sectional (A) and longitudinal (B) studies. Several publications report more than one analysis method.
doi:10.1371/journal.pone.0110300.g005

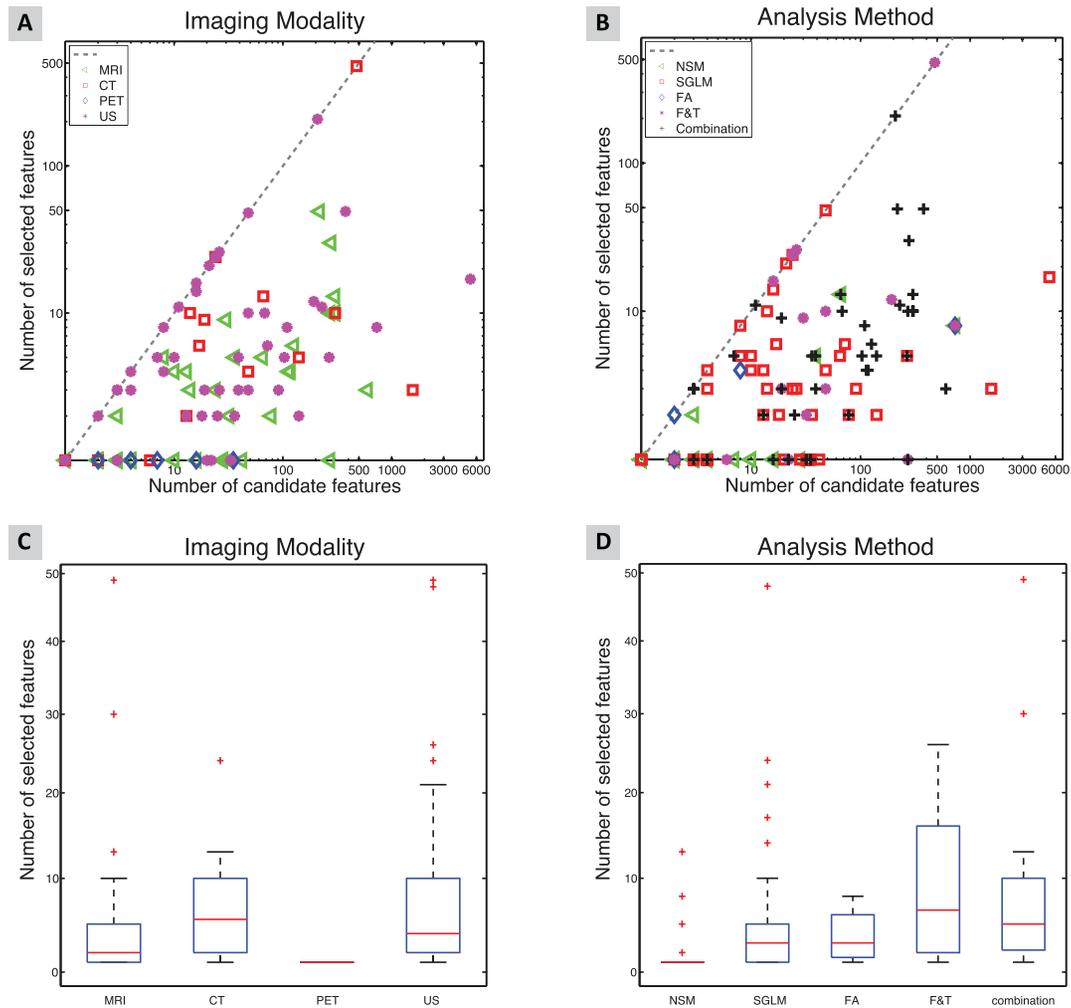


Figure 6. Number of features used in classification experiments for different imaging modalities (A) and for different analysis methods (B). Boxplot representing distribution in number selected features for imaging modality (C) and for analysis methods (D). To enhance visibility, we excluded for both boxplots two studies with large numbers of selected features. doi:10.1371/journal.pone.0110300.g006

ratio ranged from 0.46–502 (median 20) with (on average) 29% of the publications showing a tumor-feature ratio ≤ 10 .

With respect to the analysis method, publications using the F&T, or a combination of methods, had the highest risk of a tumor-feature ratio ≤ 10 , i.e. 53% and 42%, respectively. With regard to imaging modality, CT publications had the highest percentage (43%) with a tumor-feature ratio < 10 .

Using a one-way ANOVA, no significant differences were found in the performance measures between the modalities or between the analysis methods used. However, there was a negative correlation between the logarithm of the number of tumors per selected feature and the AUC ($r = -0.32$, $p < 0.05$) and the specificity ($r = -0.48$, $p < 0.05$).

Publications using statistical hypothesis testing

Of all included studies, 30% ($n = 64$) reported statistical hypothesis testing with the number of features ranging from 1–320 (median 4). Of these studies, 39% were based on MRI, 26%, on CT, 14% on PET, and 21% on US. Similarly, in 61% of the cases, data were analyzed using NSM, 12% using SGLM, 3% using FA, 6% using F&T, and 18% using a combination of these methods. The number of significant features, as reported by the

authors, ranged from 0–76 (median 1). Since multiple comparisons generally require a stronger level of evidence to be considered significant, the Holm-Bonferroni correction [20] was applied by the original research authors, or by the authors of this review paper. This correction allows direct comparison to be made of the significance levels of single and multiple comparisons. For eight papers the correction could not be performed due to missing information. After the Holm-Bonferroni correction, the number of significant features ranged from 0–6 (median 1). Figure 8 shows the number of significant features before and after the Holm-Bonferroni correction per imaging modality (A) and per analysis method (B). In 45% of the papers the number of significant features decreased after correction. Using a one-way ANOVA, no significant differences were found in the number of significant features between the modalities. With respect to the analysis method used, a one-way ANOVA established a significant difference in the number of significant features ($p < 0.018$). Publications using SGLM reported more significant features. However, after the Holm-Bonferroni correction, the numbers of significant features were similar between all analysis methods used.

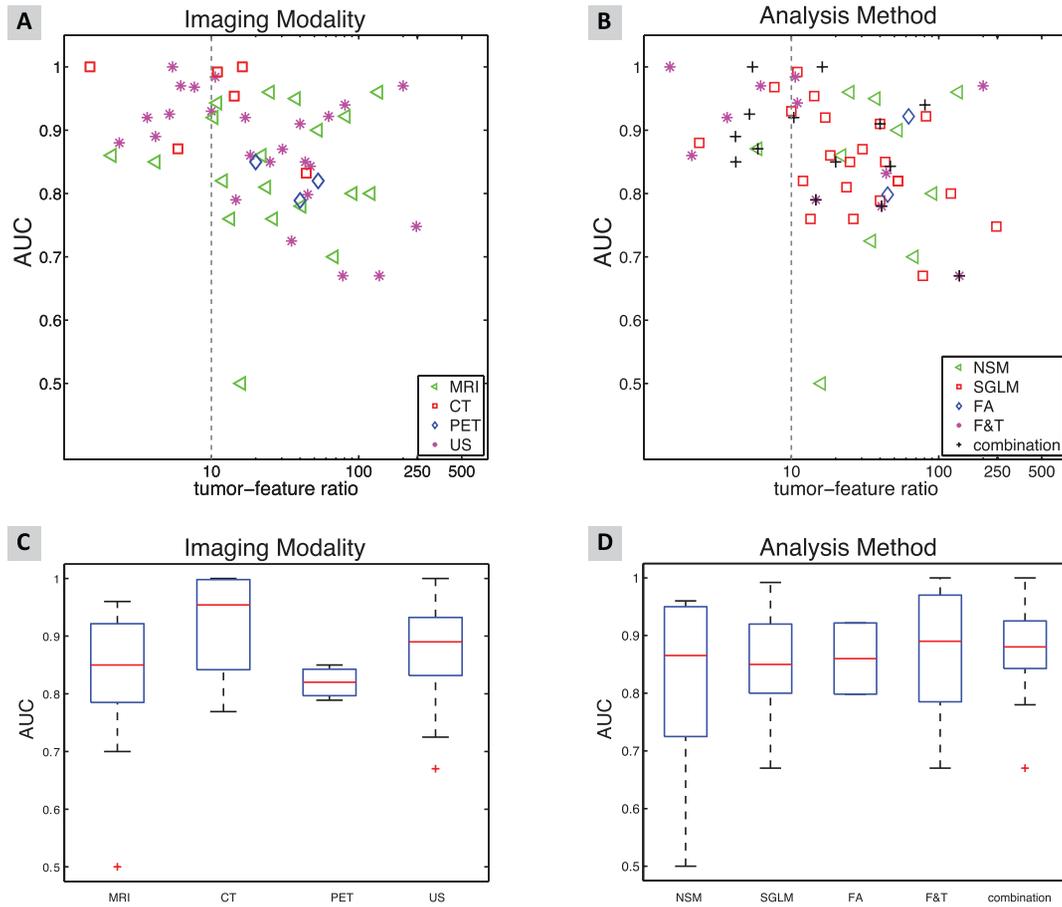


Figure 7. The AUC for different imaging modalities (A) and for different analysis methods (B) as a function of tumor-feature ratio in the classification experiments. The scatter plot shows each imaging modality and analysis method separately. Dotted line represents the ratio of 10 tumors per selected feature. Boxplot representing distribution in AUC for imaging modality (C) and for analysis methods (D).
doi:10.1371/journal.pone.0110300.g007

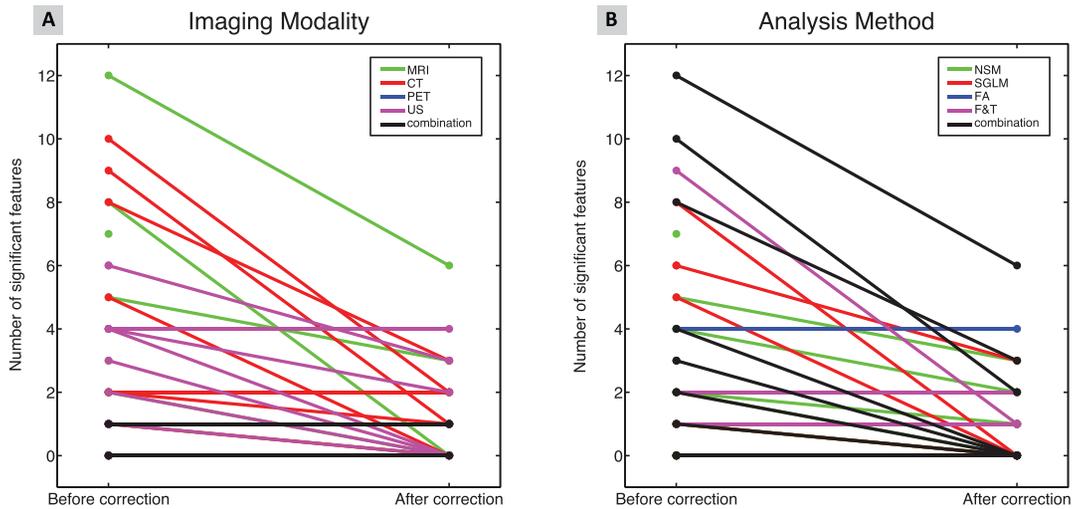


Figure 8. Number of significant features before and after Holm-Bonferroni correction in publications reporting on significance testing for all image modalities (A) and all analysis methods (B).
doi:10.1371/journal.pone.0110300.g008

Discussion

This systematic review investigated the use and performance of heterogeneity or texture quantification methods in radiological images for differentiation between tumor types, tumor grading, outcome prediction and treatment response monitoring. After a systematic literature search yielding 7351, 209 unique studies reported on heterogeneity as an imaging biomarker in tumor imaging. Since 2008, an increasing number of publications have reported on quantification of tumor heterogeneity. Since the present review is based on the existing literature, it reflects the modalities, heterogeneity analysis methods, and location of tumors that were investigated by the authors of the included studies. Because almost all of the included publications presented positive results, it should be noted that this literature probably contains an over presentation of modalities, heterogeneity analysis methods and tumor locations for which heterogeneity analysis seems to work.

Until 2006 most heterogeneity papers were based on US, whereas after 2007 there was an increase in the number of studies using MRI. During the present study period, NSM and SGLM were the most frequently used methods. Most of the papers focus on heterogeneity quantification to differentiate between tumor types, tumor grading or outcome prediction; however, the number of papers with the goal of response monitoring has recently increased. In tumor heterogeneity quantification, US is the most frequently used imaging modality for differentiation between tumor types, tumor grading and outcome prediction, and MRI is the most frequently used modality for treatment response monitoring. For monitoring of treatment response, NSM is the most frequently used method. To differentiate between tumor types and tumor grading, all methods are evenly distributed over all the modalities.

The performance of the heterogeneity features was mostly (67%) evaluated by classification experiments reporting performance measures such as accuracy, sensitivity, specificity and AUC. Papers reporting only on the results of the combination of texture features with other features were excluded from the analysis. Some authors selectively report on sensitivity without mentioning the specificity. The AUC is the preferred measure to report performance as it is more comprehensive compared to a measure based on a single threshold, such as accuracy. Only one paper reported an AUC of 0.5, all other papers reported higher values. This is most likely caused by publication bias: only the positive performance of heterogeneity features tend to reach the journals. Only 63% of the publications reporting classification results described the use of the cross-validation technique to limit the effect of over-fitting on the available data. We found no relation between the performance measures and the modality, or with the analysis method used. However, a negative correlation was found between the tumor-feature ratio and the AUC. When more tumors were available per selected feature, the AUC was lower. This correlation may be the result of overfitting of the data when fewer tumors per feature are available.

Publications using statistical hypothesis testing often did not perform a correction of the significance levels for multiple comparisons. For eight papers, due to missing information, a retrospective Holm-Bonferroni correction could not be performed by the authors. For 45% of the papers, the number of significant features decreased after the Holm-Bonferroni correction. We found no relation between the number of significant features after the Holm-Bonferroni correction and the modality or the analysis method used.

The number of prospective studies is small, i.e. only 13% of all studies. These latter studies are mainly based on MRI and report NSM features. Although the use of retrospectively collected data is necessary to develop, test and evaluate heterogeneity as a biomarker for differentiation between tumor types, tumor grading, outcome prediction and treatment response monitoring, the real test is to evaluate the performance of the developed features in a prospective study design. When using a retrospective study design, the criteria for the inclusion of cases are often not (or not clearly) described, so that the performance of the heterogeneity feature can be overestimated. Using a prospective study design, with clear inclusion criteria, the actual performance of heterogeneity features can be more reliably assessed.

Moreover, in most included studies, performance of the heterogeneity feature is evaluated without taking into account currently accepted clinical features, such as mean signal intensity, tumor size, tumor grade, or border regularity of a tumor. Some studies report only the combined classification performance of heterogeneity and clinical features. A large number of publications even use the mean signal intensity as a feature to estimate tumor heterogeneity, even though this is clearly not a heterogeneity measure (i.e., mean signal intensity does not measure intra-tumor heterogeneity). Based on these types of studies, it is not possible to evaluate the added value of heterogeneity to currently accepted clinical features. Whereas researchers are interested in the performance of the feature itself, clinicians are interested in the additional value of the feature compared with the currently available clinical biomarkers. Since the quantification of heterogeneity is usually more complex and computationally more costly than computing the mean intensity, the benefit of the added effort to characterize heterogeneity needs sufficient motivation. To enable the translation of imaging biomarkers from the research stage to clinical practice, future research should focus on studies investigating the additional value of the proposed heterogeneity biomarker compared with the established clinical markers.

In this systematic review, comparison between the performance of different methods for a certain classification task was not possible due to the large variety in the datasets used and the classification tasks posed. The search for new and optimal (combinations of) heterogeneity features would benefit from developing reliable datasets (for different classification problems) that are available to the scientific community. Large well-defined datasets are a prerequisite for objective comparison of methods.

Future studies should have a design that takes the requirements from pattern recognition into account, i.e. a balanced number of subjects and features, cross-validation, independent test datasets, and a prospective study design. Satisfying these requirements will allow more reliable evaluation of the value of heterogeneity features.

Supporting Information

Figure S1 Numbers of publications for a specific imaging modality and analysis method. The supplementary EndNote files corresponding to the records for these publications (for each cell in the matrix separately) are publically available. To download separate files just click on a cell of interest in the figure. (PDF)

Figure S2 The accuracy for different imaging modalities (A) and for different analysis methods (B) as a function of tumor-feature ratio in the classification experiments. The scatter plot shows each imaging modality and analysis method separately. Dotted line represents the ratio of 10 tumors per selected feature. Boxplot

representing distribution in AUC for imaging modality (C) and for analysis methods (D). (EPS)

Figure S3 The sensitivity for different imaging modalities (A) and for different analysis methods (B) as a function of tumor-feature ratio in the classification experiments. The scatter plot shows each imaging modality and analysis method separately. Dotted line represents the ratio of 10 tumors per selected feature. Boxplot representing distribution in AUC for imaging modality (C) and for analysis methods (D). (EPS)

Figure S4 The specificity for different imaging modalities (A) and for different analysis methods (B) as a function of tumor-feature ratio in the classification experiments. The scatter plot shows each imaging modality and analysis method separately. Dotted line represents the ratio of 10 tumors per selected feature.

References

- Fisher R, Puszta L, Swanton C (2013) Cancer heterogeneity: implications for targeted therapeutics. *Br J Cancer* 108: 479–485.
- Ng CK, Pemberton HN, Reis-Filho JS (2012) Breast cancer intratumor genetic heterogeneity: causes and implications. *Expert Rev Anticancer Ther* 12: 1021–1032.
- Brown JR, DiGiovanna MP, Killelea B, Lannin DR, Rimm DL (2014) Quantitative assessment Ki-67 score for prediction of response to neoadjuvant chemotherapy in breast cancer. *Lab Invest* 94: 98–106.
- Fasching PA, Heusinger K, Haerberle L, Niklos M, Hein A, et al. (2011) Ki67, chemotherapy response, and prognosis in breast cancer patients receiving neoadjuvant treatment. *BMC Cancer* 11: 486–498.
- Szerlip NJ, Pedraza A, Chakravarty D, Azim M, McGuire J, et al. (2012) Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. *Proc Natl Acad Sci USA* 109: 3041–3046.
- Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, et al. (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 5: 1–8.
- Hayes C, Padhani AR, Leach MO (2002) Assessing changes in tumour vascular function using dynamic contrast-enhanced magnetic resonance imaging. *NMR Biomed* 15: 154–163.
- van Rijswijk CS, Geirnaerd MJ, Hogendoorn PC, Peterse JL, van Coevorden F, et al. (2003) Dynamic contrast-enhanced MR imaging in monitoring response to isolated limb perfusion in high-grade soft tissue sarcoma: initial results. *Eur Radiol* 13: 1849–1858.
- Pickles MD, Manton DJ, Lowry M, Turnbull LW (2009) Prognostic value of pre-treatment DCE-MRI parameters in predicting disease free and overall survival for breast cancer patients undergoing neoadjuvant chemotherapy. *Eur J Radiol* 71: 498–505.
- Brizel DM, Sibley GS, Prosnitz LR, Scher RL, Dewhirst MW (1997) Tumor hypoxia adversely affects the prognosis of carcinoma of the head and neck. *Int J Radiat Oncol Biol Phys* 38: 285–289.
- Aerts HJ, Bussink J, Oyen WJ, van Elmpst W, Folgering AM, et al. (2012) Identification of residual metabolic-active areas within NSCLC tumours using a pre-radiotherapy FDG-PET-CT scan: a prospective validation. *Lung Cancer* 75: 73–76.
- Lambin P, Petit SF, Aerts HJ, van Elmpst WJ, Oberije CJ, et al. (2010) The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. *Radiother Oncol* 96: 145–152.
- PET Boost trial. Dose escalation by boosting radiation dose within the primary tumor on the basis of a pre-treatment FDG-PET-CT scan in stage IB, II and III NSCLC: a randomized Phase II trial. Available: www.clinicaltrials.gov.
- Sinha S, Lucas-Quesada FA, DeBruhl ND, Sayre J, Farria D, et al. (1997) Multifactor analysis of Gd-enhanced MR images of breast lesions. *J Magn Reson Imaging* 7: 1016–1026.
- Moher D, Liberati A, Tetzlaff J, Altman DG, Group P (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6: e1000097.
- Alic L, Veenland JV, Niessen WJ (2011) Quantification of heterogeneity as a biomarker in tumour imaging: a systematic review. Available: http://www.metaxis.com/prosperto/full_docasp?RecordID=3634 2013: 732848.
- Yang X, Knopp MV (2011) Quantifying tumor vascular heterogeneity with dynamic contrast-enhanced magnetic resonance imaging: a review. *J Biomed Biotechnol* 2011: 732848.
- Asselin MC, O'Connor JP, Boellaard R, Thacker NA, Jackson A (2012) Quantifying heterogeneity in human tumours using MRI and PET. *Eur J Cancer* 48: 447–455.
- Davnull F, Yip CS, Ljungqvist G, Selmi M, Ng F, et al. (2012) Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights Imaging* 3: 573–589.
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Statistics* 6: 65–70.
- Haralick RM, Shammugam K, Dinstein J (1973) Textural features for image classification. *IEEE Trans Syst Man Cybern* 6: 610–621.
- Amadasun M, King R (1989) Textural features corresponding to textural properties. *IEEE Trans Syst, Man Cybernet* 19: 1264–1273.
- Ojala T, Pietikainen M, Maenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 24: 971–987.
- Galloway MM (1975) Texture analysis using gray level run lengths. *Comp Graphics Image Processing* 4: 172–179.
- Ojala T, Pietikainen M, Harwood D (1996) A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition* 29: 51–59.
- Mandelbrot BB (1983) *The fractal geometry of nature*. New York: W.H. Freeman. 468 p.
- Smith TG Jr, Lange GD, Marks WB (1996) Fractal methods and results in cellular morphology—dimensions, lacunarity and multifractals. *J Neurosci Methods* 69: 123–136.
- Peleg S, Naor J, Hartley R, Avnir D (1984) Multiple resolution texture analysis and classification. *IEEE Trans Pattern Anal Mach Intell* 6: 518–523.
- Tabachnick BG, Fidell LS (2013) *Using multivariate statistics*. Boston: Pearson Education. xxxi, 983 p.
- Pekalska E, Duin RPW (2005) *The dissimilarity representation for pattern recognition: foundations and applications*. Hackensack, N.J.: World Scientific. xxvi, 607 p.
- Young TY, Calvert TW (1974) *Classification, estimation, and pattern recognition*: American Elsevier Pub. Co. 366 p.
- Acharya UR, Faust O, Sree SV, Molinari F, Garberoglio R, et al. (2011) Cost-effective and non-invasive automated benign and malignant thyroid lesion classification in 3D contrast-enhanced ultrasound using combination of wavelets and textures: a class of ThyroScan algorithms. *Technol Cancer Res Treat* 10: 371–380.
- Acharya UR, Faust O, Sree SV, Molinari F, Suri JS (2012) ThyroScreen system: high resolution ultrasound thyroid image characterization into benign and malignant classes using novel combination of texture and discrete wavelet transform. *Comput Meth Progr Biomed* 107: 233–241.
- Chang RF, Wu WJ, Moon WK, Chen DR (2003) Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis. *Ultrasound Med Biol* 29: 679–686.
- Chang RF, Wu WJ, Moon WK, Chou YH, Chen DR (2003) Support vector machines for diagnosis of breast tumors on US images. *Acad Radiol* 10: 189–197.
- Chen D, Chang RF, Huang YL (2000) Breast cancer diagnosis using self-organizing map for sonography. *Ultrasound Med Biol* 26: 405–411.
- Chen DR, Chang RF, Huang YL (1999) Computer-aided diagnosis applied to US of solid breast nodules by using neural networks. *Radiology* 213: 407–412.
- Chen DR, Kuo WJ, Chang RF, Moon WK, Lee CC (2002) Use of the bootstrap technique with small training sets for computer-aided diagnosis in breast ultrasound. *Ultrasound Med Biol* 28: 897–902.
- Chen SJ, Cheng KS, Dai YC, Sun YN, Chen YT, et al. (2005) Quantitatively characterizing the textural features of sonographic images for breast cancer with histopathologic correlation. *J Ultrasound Med* 24: 651–661.

40. Chen W, Giger ML, Bick U, Newstead GM (2006) Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI. *Med Phys* 33: 2878–2887.
41. Ganeshan B, Abaleke S, Young RC, Chatwin CR, Miles KA (2010) Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer Imaging* 10: 137–143.
42. Georgiadis P, Cavouras D, Kalatzis I, Daskalakis A, Kagadis GC, et al. (2008) Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features. *Comput Meth Programs Biomed* 89: 24–32.
43. Georgiadis P, Kostopoulos S, Cavouras D, Glotsos D, Kalatzis I, et al. (2011) Quantitative combination of volumetric MR imaging and MR spectroscopy data for the discrimination of meningiomas from metastatic brain tumors by means of pattern recognition. *Magn Reson Imaging* 29: 525–535.
44. Harrison L, Dastidar P, Eskola H, Jarvenpaa R, Pertovaara H, et al. (2008) Texture analysis on MRI images of non-Hodgkin lymphoma. *Comput Biol Med* 38: 519–524.
45. Kido S, Kuriyama K, Higashiyama M, Kasugai T, Kuroda C (2002) Fractal analysis of small peripheral pulmonary nodules in thin-section CT: evaluation of the lung-nodule interfaces. *J Comput Assist Tomogr* 26: 573–578.
46. Klein HM, Klose KC, Eisele T, Brenner M, Ameling W, et al. (1993) [The diagnosis of focal liver lesions by the texture analysis of dynamic computed tomograms]. *Rofo* 159: 10–15.
47. McNitt-Gray MF, Hart EM, Wyckoff N, Sayre JW, Goldin JG, et al. (1999) A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: preliminary results. *Med Phys* 26: 880–888.
48. Ng F, Kozarski R, Ganeshan B, Goh V (2013) Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? *Eur J Radiol* 82: 342–348.
49. O'Sullivan F, Roy S, Eary J (2003) A statistical measure of tissue heterogeneity with application to 3D PET sarcoma data. *Biostatistics* 4: 433–448.
50. Sun T, Wang J, Li X, Lv P, Liu F, et al. (2013) Comparative evaluation of support vector machines for computer aided diagnosis of lung cancer in CT based on a multi-dimensional data set. *Comput Meth Programs Biomed* 111: 519–524.
51. Thijssen JM, Verbeek AM, Romijn RL, de Wolff-Rouendaal D, Oosterhuis JA (1991) Echographic differentiation of histological types of intraocular melanoma. *Ultrasound Med Biol* 17: 127–138.
52. Way TW, Hadjiiski LM, Sahiner B, Chan HP, Cascade PN, et al. (2006) Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours. *Med Phys* 33: 2323–2337.
53. Wu WJ, Moon WK (2008) Ultrasound breast tumor image computer-aided diagnosis with texture and morphological features. *Acad Radiol* 15: 873–880.
54. Chen DR, Chang RF, Huang YL, Chou YH, Tiu CM, et al. (2000) Texture analysis of breast tumors on sonograms. *Semin Ultrasound CT MR* 21: 308–316.
55. Chen DR, Chang RF, Kuo WJ, Chen MC, Huang YL (2002) Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. *Ultrasound Med Biol* 28: 1301–1310.
56. Chen DR, Huang YL, Lin SH (2011) Computer-aided diagnosis with textural features for breast lesions in sonograms. *Comput Med Imaging Graph* 35: 220–226.
57. Chen DR, Liang WM, Kuo HW, Chang RF (1999) Computerized quantitative assessment of sonomammographic homogeneity of fibroadenoma and breast carcinoma. *J of Med Ultrasound* 7: 157–162.
58. Chen EL, Chung YN, Chung PC, Tsai HM, Huang YS (2001) Using a fuzzy engine and complete set of features for hepatic diseases diagnosis: Integrating contrast and non-contrast CT images. *Biomed Eng - Applications, Basis and Communications* 13: 159–167.
59. Chen SJ, Chang CY, Chang KY, Tzeng JE, Chen YT, et al. (2010) Classification of the thyroid nodules based on characteristic sonographic textural feature and correlated histopathology using hierarchical support vector machines. *Ultrasound Med Biol* 36: 2018–2026.
60. Chen SJ, Cheng KS, Dai YC, Sun YN, Chen YT, et al. (2005) The representations of sonographic image texture for breast cancer using co-occurrence matrix. *J Med and Biol Eng* 25: 193–199.
61. Chen SJ, Lin CH, Chang CY, Chang KY, Ho HC, et al. (2012) Characterizing the major sonographic textural difference between metastatic and common benign lymph nodes using support vector machine with histopathologic correlation. *Clin Imaging* 36: 353–359 e352.
62. Chen SJ, Yu SN, Tzeng JE, Chen YT, Chang KY, et al. (2009) Characterization of the major histopathological components of thyroid nodules using sonographic textural features for clinical diagnosis and management. *Ultrasound Med Biol* 35: 201–208.
63. Chen W, Giger ML, Li H, Bick U, Newstead GM (2007) Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. *Magn Reson Med* 58: 562–571.
64. Chen WM, Chang RF, Kuo SJ, Chang CS, Moon WK, et al. (2005) 3-D ultrasound texture classification using run difference matrix. *Ultrasound Med Biol* 31: 763–770.
65. Chikui T, Tokumori K, Yoshiura K, Oobu K, Nakamura S, et al. (2005) Sonographic texture characterization of salivary gland tumors by fractal analyses. *Ultrasound Med Biol* 31: 1297–1304.
66. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, et al. (2013) Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nuc Med* 54: 19–26.
67. Cui C, Cai H, Liu L, Li L, Tian H, et al. (2011) Quantitative analysis and prediction of regional lymph node status in rectal cancer based on computed tomography imaging. *Eur Radiol* 21: 2318–2325.
68. Cui J, Sahiner B, Chan HP, Nees A, Paramagul C, et al. (2009) A new automated method for the segmentation and characterization of breast masses on ultrasound images. *Med Phys* 36: 1553–1565.
69. de Langen AJ, van den Boogaart V, Lubberink M, Backes WH, Marcus JT, et al. (2011) Monitoring response to antiangiogenic therapy in non-small cell lung cancer using imaging markers derived from PET and dynamic contrast-enhanced MRI. *J Nucl Med* 52: 48–55.
70. de Lussanet QG, Backes WH, Griffioen AW, Padhani AR, Baeten CI, et al. (2005) Dynamic contrast-enhanced magnetic resonance imaging of radiation therapy-induced microcirculation changes in rectal cancer. *Int J Radiat Oncol Biol Phys* 63: 1309–1315.
71. Ding J, Cheng H, Ning C, Huang J, Zhang Y (2011) Quantitative measurement for thyroid cancer characterization based on elastography. *J Ultrasound Med* 30: 1259–1266.
72. Dominietto M, Lehmann S, Keist R, Rudin M (2012) Pattern analysis accounts for heterogeneity observed in MRI studies of tumor angiogenesis. *Magn Reson Med* 70: 1481–1490.
73. Dong X, Xing L, Wu P, Fu Z, Wan H, et al. (2013) Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor 18F-fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage. *Nucl Med Commun* 34: 40–46.
74. Donohue KD, Forsberg F, Piccoli CV, Goldberg BB (1999) Analysis and classification of tissue with scatterer structure templates. *IEEE Trans Ultrason Ferroelectr Freq Control* 46: 300–310.
75. Donohue KD, Huang L, Burks T, Forsberg F, Piccoli CW (2001) Tissue classification with generalized spectrum parameters. *Ultrasound Med Biol* 27: 1505–1514.
76. Downey K, Riches SF, Morgan VA, Giles SL, Attygalle AD, et al. (2013) Relationship between imaging biomarkers of stage I cervical cancer and poor-prognosis histologic features: quantitative histogram analysis of diffusion-weighted MR images. *Am J Roentgenol* 200: 314–320.
77. Drabycz S, Roldan G, de Robles P, Adler D, McIntyre JB, et al. (2010) An analysis of image texture, tumor location, and MGMT promoter methylation in glioblastoma using magnetic resonance imaging. *Neuroimage* 49: 1398–1405.
78. Dumrongpisutikul N, Intrapirromkul J, Yousem DM (2012) Distinguishing between germinomas and pineal cell tumors on MR imaging. *AJNR Am J Neuroradiol* 33: 550–555.
79. Eary JF, O'Sullivan F, O'Sullivan J, Conrad EU (2008) Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome. *J Nucl Med* 49: 1973–1979.
80. Eliat PA, Lechaux D, Gervais A, Rioux-Leclerc N, Franconi F, et al. (2001) Is magnetic resonance imaging texture analysis a useful tool for cell therapy in vivo monitoring? *Anticancer Res* 21: 3857–3860.
81. Eliat PA, Olivie D, Saikali S, Carsin B, Saint-Jalmes H, et al. (2012) Can dynamic contrast-enhanced magnetic resonance imaging combined with texture analysis differentiate malignant glioneuronal tumors from other glioblastoma? *Neur Res Int* 2012: 1–7.
82. Emblem KE, Nedregaard B, Nome T, Due-Tonnessen P, Hald JK, et al. (2008) Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. *Radiology* 247: 808–817.
83. Engelbrecht MR, Hitge-Boetes C, Coolen J, Thijssen JM, Makkus AC, et al. (1998) Follow-up of Wilms' tumour during pre-operative chemotherapy by qualitative and quantitative sonography. *Eur J Ultrasound* 8: 157–165.
84. Farace P, Galie M, Merigo F, Daducci A, Calderan L, et al. (2009) Inhibition of tyrosine kinase receptors by SU6668 promotes abnormal stromal development at the periphery of carcinomas. *Br J Cancer* 100: 1575–1580.
85. Faschingbauer F, Beckmann MW, Weyert Goecke T, Renner S, Haberer L, et al. (2013) Automatic texture-based analysis in ultrasound imaging of ovarian masses. *Ultraschall Med* 34: 145–150.
86. Fetit AE, Novak J, Rodriguez D, Auer DP, Clark CA, et al. (2013) MRI texture analysis in paediatric oncology: a preliminary study. *Stud Health Technol Inform* 190: 169–171.
87. Fruehwald-Pallamar J, Czerny C, Holzer-Fruehwald L, Nemecek SF, Mueller-Mang C, et al. (2013) Texture-based and diffusion-weighted discrimination of parotid gland lesions on MR images at 3.0 Tesla. *NMR Biomed* 26: 1372–1379.
88. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K (2012) Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *Eur Radiol* 22: 796–802.
89. Ganeshan B, Skogen K, Pressney I, Coutroubis D, Miles K (2012) Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol* 67: 157–164.

90. Garra BS, Krasner BH, Horii SC, Ascher S, Mun SK, et al. (1993) Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. *Ultrasound Imaging* 15: 267–285.
91. Gensure RH, Foran DJ, Lee VM, Gendel VM, Jabbour SK, et al. (2012) Evaluation of hepatic tumor response to yttrium-90 radioembolization therapy using texture signatures generated from contrast-enhanced CT images. *Acad Radiol* 19: 1201–1207.
92. Georgiadis P, Cavouras D, Kalatzis I, Glotsos D, Athanasiadis E, et al. (2009) Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain MRI by volumetric textural features and ensemble pattern recognition methods. *Magn Reson Imaging* 27: 120–130.
93. Gibbs P, Turnbull LW (2003) Textural analysis of contrast-enhanced MR images of the breast. *Magn Reson Med* 50: 92–98.
94. Giger ML, Al-Hallaq H, Huo Z, Moran C, Wolverton DE, et al. (1999) Computerized analysis of lesions in US images of the breast. *Acad Radiol* 6: 665–674.
95. Gletsos M, Mougiakakou SG, Matsopoulos GK, Nikita KS, Nikita AS, et al. (2003) A computer-aided diagnostic system to characterize CT focal liver lesions: design and optimization of a neural network classifier. *IEEE Trans Inf Technol Biomed* 7: 153–162.
96. Glotsos D, Kalatzis I, Theodorakis P, Georgiadis P, Daskalakis A, et al. (2010) A multi-classifier system for the characterization of normal, infectious, and cancerous prostate tissues employing transrectal ultrasound images. *Comput Meth Programs Biomed* 97: 53–61.
97. Goh V, Ganeshan B, Nathan P, Jutla JK, Vinayan A, et al. (2011) Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker. *Radiology* 261: 165–171.
98. Goldberg V, Manduca A, Ewert DL, Gisvold JJ, Greenleaf JF (1992) Improvement in specificity of ultrasonography for diagnosis of breast tumors by means of artificial intelligence. *Med Phys* 19: 1475–1481.
99. Gomez W, Pereira WC, Infantosi AF (2012) Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound. *IEEE Trans Med Imaging* 31: 1889–1899.
100. Haney CR, Fan X, Markiewicz E, Mustafi D, Karczmar GS, et al. (2013) Monitoring anti-angiogenic therapy in colorectal cancer murine model using dynamic contrast-enhanced MRI: comparing pixel-by-pixel with region of interest analysis. *Technol Cancer Res Treat* 12: 71–78.
101. Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D (2013) Robustness of intratumour (18)F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging* 40: 1662–1671.
102. Herts BR, Coll DM, Novick AC, Obuchowski N, Linnell G, et al. (2002) Enhancement characteristics of papillary renal neoplasms revealed on triphasic helical CT of the kidneys. *AJR Am J Roentgenol* 178: 367–372.
103. Hirano M, Satake H, Ishigaki S, Ikeda M, Kawai H, et al. (2012) Diffusion-weighted imaging of breast masses: comparison of diagnostic performance using various apparent diffusion coefficient parameters. *AJR Am J Roentgenol* 198: 717–722.
104. Hirning T, Zuna I, Schlaps D, Lorenz D, Meybier H, et al. (1989) Quantification and classification of echographic findings in the thyroid gland by computerized B-mode texture analysis. *Eur J Radiol* 9: 244–247.
105. Holli K, Laaperi AL, Harrison L, Luukkaala T, Toivonen T, et al. (2010) Characterization of breast cancer types by texture analysis of magnetic resonance images. *Acad Radiol* 17: 135–141.
106. Horsch K, Giger ML, Venta LA, Vyborny CJ (2002) Computerized diagnosis of breast lesions on ultrasound. *Med Phys* 29: 157–164.
107. Huang B, Chan T, Kwong DL, Chan WK, Khong PL (2012) Nasopharyngeal carcinoma: investigation of intratumoral heterogeneity with FDG PET/CT. *AJR Am J Roentgenol* 199: 169–174.
108. Huang YL, Chen JH, Shen WC (2006) Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images. *Acad Radiol* 13: 713–720.
109. Huang YL, Kuo SJ, Chang CS, Liu YK, Moon WK, et al. (2005) Image retrieval with principal component analysis for breast cancer diagnosis on various ultrasonic systems. *Ultrasound Obstet Gynecol* 26: 558–566.
110. Huang Z, Mayr NA, Lo SS, Grecula JC, Wang JZ, et al. (2012) Characterizing at-risk voxels by using perfusion magnetic resonance imaging for cervical cancer during radiotherapy. *J Cancer Sci Ther* 4: 254–259.
111. Huber S, Danes J, Zuna I, Teubner J, Medl M, et al. (2000) Relevance of sonographic B-mode criteria and computer-aided ultrasonic tissue characterization in differential/diagnosis of solid breast masses. *Ultrasound Med Biol* 26: 1243–1252.
112. Iakovidis DK, Keramidas EG, Maroulis D (2010) Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns. *Artif Intell Med* 50: 33–41.
113. Issa B, Buckley DL, Turnbull LW (1999) Heterogeneity analysis of Gd-DTPA uptake: improvement in breast lesion differentiation. *J Comput Assist Tomogr* 23: 615–621.
114. Jansen JF, Schoder H, Lee NY, Stambuk HE, Wang Y, et al. (2012) Tumor metabolism and perfusion in head and neck squamous cell carcinoma: pretreatment multimodality imaging with 1H magnetic resonance spectroscopy, dynamic contrast-enhanced MRI, and [18F]FDG-PET. *Int J Radiat Oncol Biol Phys* 82: 299–307.
115. Jung SC, Cho JY, Kim SH (2012) Subtype differentiation of small renal cell carcinomas on three-phase MDCT: usefulness of the measurement of degree and heterogeneity of enhancement. *Acta Radiol* 53: 112–118.
116. Juntu J, Sibjers J, De Backer S, Rajan J, Van Dyck D (2010) Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. *J Magn Reson Imaging* 31: 680–689.
117. Karahaliou A, Vassiou K, Arikidis NS, Skiadopoulos S, Kanavou T, et al. (2010) Assessing heterogeneity of lesion enhancement kinetics in dynamic contrast-enhanced MRI for breast cancer diagnosis. *Br J Radiol* 83: 296–309.
118. Kidd EA, Grigsby PW (2008) Intratumoral metabolic heterogeneity of cervical cancer. *Clin Cancer Res* 14: 5236–5241.
119. Kido S, Kuriyama K, Higashiyama M, Kasugai T, Kuroda C (2003) Fractal analysis of internal and peripheral textures of small peripheral bronchogenic carcinomas in thin-section computed tomography: comparison of bronchioalveolar cell carcinomas with nonbronchioalveolar cell carcinomas. *J Comput Assist Tomogr* 27: 56–61.
120. Kim DY, Kim JH, Noh SM, Park JW (2003) Pulmonary nodule detection using chest CT images. *Acta Radiol* 44: 252–257.
121. Kim KG, Cho SW, Min SJ, Kim JH, Min BG, et al. (2005) Computerized scheme for assessing ultrasonographic features of breast masses. *Acad Radiol* 12: 58–66.
122. Kim KG, Kim JH, Min BG (2001) Comparative analysis of texture characteristics of malignant and benign tumors in breast ultrasonograms. *J Digit Imaging* 14: 208–210.
123. Kjaer L, Ring P, Thomsen C, Henriksen O (1995) Texture analysis in quantitative MR imaging. Tissue characterisation of normal brain and intracranial tumours at 1.5 T. *Acta Radiol* 36: 127–135.
124. Klein HM, Eisele T, Klose KC, Stauss I, Brenner M, et al. (1996) Pattern recognition system for focal liver lesions using “crisp” and “fuzzy” classifiers. *Invest Radiol* 31: 6–10.
125. Kratzik C, Schuster E, Hainz A, Kuber W, Lunglmayr G (1988) Texture analysis—a new method of differentiating prostatic carcinoma from prostatic hypertrophy. *Urol Res* 16: 395–397.
126. Kuntz C, Glaser F, Zuna I, Buhr HJ, Herfarth C (1994) Endorectal ultrasound and computerized B-scan texture analysis to assess sessile adenoma and small rectal carcinoma. *Endoskopie Heute* 7: 173–178.
127. Kuo WJ, Chang RF, Lee CC, Moon WK, Chen DR (2002) Retrieval technique for the diagnosis of solid breast tumors on sonogram. *Ultrasound Med Biol* 28: 903–909.
128. Kuo WJ, Chang RF, Moon WK, Lee CC, Chen DR (2002) Computer-aided diagnosis of breast tumors with different US systems. *Acad Radiol* 9: 793–799.
129. Kurki T, Lundbom N, Kalimo H, Valtonen S (1995) MR classification of brain gliomas: value of magnetization transfer and conventional imaging. *Magn Reson Imaging* 13: 501–511.
130. Lai YC, Huang YS, Wang DW, Tiu CM, Chou YH, et al. (2013) Computer-aided diagnosis for 3-d power Doppler breast ultrasound. *Ultrasound Med Biol* 39: 555–567.
131. Larkin TJ, Canuto HC, Kettunen MI, Booth TC, Hu DE, et al. (2013) Analysis of image heterogeneity using 2D Minkowski functionals detects tumor responses to treatment. *Magn Reson Med* 7.
132. Lee CC, Shih CY (2010) Learning patterns of liver masses using improved RBF networks. *Biomedical Engineering - Applications, Basis and Communications* 22: 137–147.
133. Lefebvre F, Meunier M, Thibault F, Laugier P, Berger G (2000) Computerized ultrasound B-scan characterization of breast nodules. *Ultrasound Med Biol* 26: 1421–1428.
134. Li X, Lu Y, Pirzkall A, McKnight T, Nelson SJ (2002) Analysis of the spatial characteristics of metabolic abnormalities in newly diagnosed glioma patients. *J Magn Reson Imaging* 16: 229–237.
135. Liao YY, Tsui PH, Li CH, Chang KJ, Kuo WH, et al. (2011) Classification of scattering media within benign and malignant breast tumors based on ultrasound texture-feature-based and Nakagami-parameter images. *Med Phys* 38: 2198–2207.
136. Liu F, Kornecki A, Shmullovich O, Gelman N (2011) Optimization of time-to-peak analysis for differentiating malignant and benign breast lesions with dynamic contrast-enhanced MRI. *Acad Radiol* 18: 694–704.
137. Liu Y, Cheng HD, Huang JH, Zhang YT, Tang XL, et al. (2012) Computer aided diagnosis system for breast cancer based on color Doppler flow imaging. *J Med Syst* 36: 3975–3982.
138. Liu YH, Muftah M, Das T, Bai L, Robson K, et al. (2012) Classification of MR tumor images based on Gabor wavelet analysis. *J Med Biol Eng* 32: 22–28.
139. Loren DE, Seghal CM, Ginsberg GG, Kochman ML (2002) Computer-assisted analysis of lymph nodes detected by EUS in patients with esophageal carcinoma. *Gastrointest Endosc* 56: 742–746.
140. Ma JH, Kim HS, Rim NJ, Kim SH, Cho KG (2010) Differentiation among glioblastoma multiforme, solitary metastatic tumor, and lymphoma using whole-tumor histogram analysis of the normalized cerebral blood volume in enhancing and enhancing lesions. *Am J Neuroradiol* 31: 1699–1706.
141. Maruyama H, Takahashi M, Sekimoto T, Kamesaki H, Shimada T, et al. (2012) Heterogeneity of microbubble accumulation: a novel approach to discriminate between well-differentiated hepatocellular carcinomas and regenerative nodules. *Ultrasound Med Biol* 38: 383–388.

142. Mattonen SA, Palma DA, Haasbeek CJ, Senan S, Ward AD (2013) Distinguishing radiation fibrosis from tumour recurrence after stereotactic ablative radiotherapy (SABR) for lung cancer: a quantitative analysis of CT density changes. *Acta Oncol* 52: 910–918.
143. Mayerhoefer ME, Breitenscher M, Amann G, Dominkus M (2008) Are signal intensity and homogeneity useful parameters for distinguishing between benign and malignant soft tissue masses on MR images? Objective evaluation by means of texture analysis. *Magn Reson Imaging* 26: 1316–1322.
144. Mayerhoefer ME, Schima W, Trattnig S, Pinker K, Berger-Kulemann V, et al. (2010) Texture-based classification of focal liver lesions on MRI at 3.0 Tesla: a feasibility study in cysts and hemangiomas. *J Magn Reson Imaging* 32: 352–359.
145. Mayr NA, Yuh WT, Arnholt JC, Ehrhardt JC, Sorosky JI, et al. (2000) Pixel analysis of MR perfusion imaging in predicting radiation therapy outcome in cervical cancer. *J Magn Reson Imaging* 12: 1027–1033.
146. McLaren CE, Chen WP, Nie K, Su MY (2009) Prediction of malignant breast lesions from MRI features: a comparison of artificial neural network and logistic regression techniques. *Acad Radiol* 16: 842–851.
147. McNitt-Gray MF, Wyckoff N, Sayre JW, Goldin JG, Aberle DR (1999) The effects of co-occurrence matrix based texture parameters on the classification of solitary pulmonary nodules imaged on computed tomography. *Comput Med Imaging Graph* 23: 339–348.
148. Meinel LA, Stolpen AH, Berbaum KS, Fajardo LL, Reinhardt JM (2007) Breast MRI lesion classification: improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. *J Magn Reson Imaging* 25: 89–95.
149. Miles KA, Ganeshan B, Griffiths MR, Young RC, Chatwin CR (2009) Colorectal cancer: texture analysis of portal phase hepatic CT images as a potential marker of survival. *Radiology* 250: 444–452.
150. Mitrea D, Mitrea P, Nedevschi S, Badea R, Lupșor M, et al. (2012) Abdominal tumor characterization and recognition using superior-order cooccurrence matrices, based on ultrasound images. *Comput Math Methods Med* 2012: 348135.
151. Mittal D, Kumar V, Saxena SC, Khandelwal N, Kalra N (2011) Neural network based focal liver lesion diagnosis using ultrasound images. *Comput Med Imaging Graph* 35: 315–323.
152. Moon WK, Shen YW, Huang CS, Chiang LR, Chang RF (2011) Computer-aided diagnosis for the classification of breast masses in automated whole breast ultrasound images. *Ultrasound Med Biol* 37: 539–548.
153. Mougialakou SG, Valavanis IK, Nikita A, Nikita KS (2007) Differential diagnosis of CT focal liver lesions using texture features, feature selection and ensemble driven classifiers. *Artif Intell Med* 41: 25–37.
154. Mouthuy N, Cosnard G, Abarca-Quinones J, Michoux N (2012) Multiparametric magnetic resonance imaging to differentiate high-grade gliomas and brain metastases. *J Neuroradiol* 39: 301–307.
155. Mussurakis S, Gibbs P, Horsman A (1998) Peripheral enhancement and spatial contrast uptake heterogeneity of primary breast tumours: quantitative assessment with dynamic MRI. *J Comput Assist Tomogr* 22: 35–46.
156. Nagarajan MB, Huber MB, Schlossbauer T, Leinsinger G, Krol A, et al. (2013) Classification of Small Lesions in Breast MRI: Evaluating The Role of Dynamically Extracted Texture Features Through Feature Selection. *J Med Biol Eng* 33: 59–68.
157. Negendank WG, al-Katib AM, Karanes C, Smith MR (1990) Lymphomas: MR imaging contrast characteristics with clinical-pathologic correlations. *Radiology* 177: 209–216.
158. Newell D, Nie K, Chen JH, Hsu CC, Yu HJ, et al. (2010) Selection of diagnostic features on breast MRI to differentiate between malignant and benign lesions using computer-aided diagnosis: differences in lesions presenting as mass and non-mass-like enhancement. *Eur Radiol* 20: 771–781.
159. Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V (2013) Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology* 266: 177–184.
160. Nguyen P, Bashirzadeh F, Hundloe J, Salvado O, Dowson N, et al. (2012) Optical differentiation between malignant and benign lymphadenopathy by grey scale texture analysis of endobronchial ultrasound convex probe images. *Chest* 141: 709–715.
161. Nguyen VX, Nguyen CC, Li B, Das A (2010) Digital image analysis is a useful adjunct to endoscopic ultrasonographic diagnosis of subepithelial lesions of the gastrointestinal tract. *J Ultrasound Med* 29: 1345–1351.
162. Nie K, Chen JH, Yu HJ, Chu Y, Nalcioğlu O, et al. (2008) Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI. *Acad Radiol* 15: 1513–1525.
163. O'Connor JP, Rose CJ, Jackson A, Watson Y, Cheung S, et al. (2011) DCE-MRI biomarkers of tumour heterogeneity predict CRC liver metastasis shrinkage following bevacizumab and FOLFOX-6. *Br J Cancer* 105: 139–145.
164. O'Sullivan F, Wolsztynski E, O'Sullivan J, Richards T, Conrad EU, et al. (2011) A statistical modeling approach to the analysis of spatial patterns of FDG-PET uptake in human sarcoma. *IEEE Trans Med Imaging* 30: 2059–2071.
165. Padma A, Sukanesh R (2013) Combined texture feature analysis of segmentation and classification of benign and malignant tumour CT slices. *J Med Eng Technol* 37: 1–9.
166. Pang Y, Li L, Hu W, Peng Y, Liu L, et al. (2012) Computerized segmentation and characterization of breast lesions in dynamic contrast-enhanced MR images using fuzzy c-means clustering and snake algorithm. *Comput Math Methods Med* 2012: 634907.
167. Peeters F, Rommel D, Abarca-Quinones J, Gregoire V, Duprez T (2012) Early (72-hour) detection of radiotherapy-induced changes in an experimental tumor model using diffusion-weighted imaging, diffusion tensor imaging, and Q-space imaging parameters: a comparative study. *J Magn Reson Imaging* 35: 409–417.
168. Peng SL, Chen CF, Liu HL, Lui CC, Huang YJ, et al. (2012) Analysis of parametric histogram from dynamic contrast-enhanced MRI: application in evaluating brain tumor response to radiotherapy. *NMR Biomed* 26: 443–451.
169. Piliouras N, Kalatzis I, Dimitropoulos N, Cavouras D (2004) Development of the cubic least squares mapping linear-kernel support vector machine classifier for improving the characterization of breast lesions on ultrasound. *Comput Med Imaging Graph* 28: 247–255.
170. Plant RL (1998) Image analysis of benign and malignant neck masses. *Ann Otol Rhinol Laryngol* 107: 689–696.
171. Preim U, Glasser S, Preim B, Fischbach F, Ricke J (2012) Computer-aided diagnosis in breast DCE-MRI—quantification of the heterogeneity of breast lesions. *Eur J Radiol* 81: 1532–1538.
172. Prescott JW, Zhang D, Wang JZ, Mayr NA, Yuh WT, et al. (2010) Temporal analysis of tumor heterogeneity and volume for cervical cancer treatment outcome prediction: preliminary evaluation. *J Digit Imaging* 23: 342–357.
173. Raja JV, Khan M, Ramachandra VK, Al-Kadi O (2012) Texture analysis of CT images in the characterization of oral cancers involving buccal mucosa. *Dentomaxillofac Radiol* 41: 475–480.
174. Rakoczy M, McGaughey D, Korenberg MJ, Levman J, Martel AL (2013) Feature selection in computer-aided breast cancer diagnosis via dynamic contrast-enhanced magnetic resonance images. *J Digit Imaging* 26: 198–208.
175. Ravanelli M, Farina D, Morassi M, Roca E, Cavalleri G, et al. (2013) Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy. *Eur Radiol* 23: 3450–3454.
176. Rehn S, Sperber GO, Nyman R, Glimelius B, Hagberg H, et al. (1993) Quantification of inhomogeneities in malignancy grading of non-Hodgkin lymphoma with MR imaging. *Acta Radiol* 34: 3–9.
177. Rehn SM, Rodriguez M, Bergstrom SR, Nyman RS, Glimelius BL (1997) Tumour inhomogeneities on magnetic resonance imaging, a new factor with prognostic information in non-Hodgkin's lymphomas. *Leuk Lymphoma* 24: 501–511.
178. Retico A, Fantacci ME, Gori I, Kasae P, Golosio B, et al. (2009) Pleural nodule identification in low-dose and thin-slice lung computed tomography. *Comput Biol Med* 39: 1137–1144.
179. Revert Ventura AJ, Sanz Requena R, Marti-Bonmati L, Pallardo Y, Jornt J, et al. (2012) The heterogeneity of blood flow on magnetic resonance imaging: a biomarker for grading cerebral astrocytomas. *Radiologia* 56.
180. Rokita E, Taton G, Sierzeza M, Kulig J, Urbanik A (2009) Quantitative analysis of 3D US images in the relationship with liver lesion diagnosis. *Polish J Radiol* 74: 28–34.
181. Rose CJ, Mills SJ, O'Connor JP, Buonaccorsi GA, Roberts C, et al. (2009) Quantifying spatial heterogeneity in dynamic contrast-enhanced MRI parameter maps. *Magn Reson Med* 62: 488–499.
182. Sachdeva J, Kumar V, Gupta I, Khandelwal N, Ahuja CK (2012) A dual neural network ensemble approach for multiclas brain tumor classification. *Int J Numer Method Biomed Eng* 28: 1107–1120.
183. Sadeghi-Naini A, Falou O, Tadayyon H, Al-Mahrouki A, Tran W, et al. (2013) Conventional frequency ultrasonic biomarkers of cancer treatment response in vivo. *Transl Oncol* 6: 234–243.
184. Sahiner B, Chan HP, Roubidoux MA, Helvie MA, Hadjiiski LM, et al. (2004) Computerized characterization of breast masses on three-dimensional ultrasound volumes. *Med Phys* 31: 744–754.
185. Sarty GE, Kendall EJ, Loewy J, Dhir A, Olatunbosun OA, et al. (2004) Magnetic resonance diffusion imaging of ovarian masses: a first experience with 12 cases. *MAGMA* 16: 182–193.
186. Sasikala M, Kumaravel N (2008) A wavelet-based optimal texture feature set for classification of brain tumours. *J Med Eng Technol* 32: 198–205.
187. Scheipers U, Siebers S, Gottwald F, Ashfaq M, Bozzato A, et al. (2005) Sonohistology for the computerized differentiation of parotid gland tumors. *Ultrasound Med Biol* 31: 1287–1296.
188. Schulte M, von Baer A, Schultheiss M, Scheil-Bertram S (2010) [Classification of solid soft tissue tumours by ultrasonography]. *Ultraschall Med* 31: 182–190.
189. Selvan S, Kavitha M, Devi SS, Suresh S (2012) Fuzzy-based classification of breast lesions using ultrasound echography and elastography. *Ultrasound Q* 28: 159–167.
190. Shah SK, McNitt-Gray MF, Rogers SR, Goldin JG, Suh RD, et al. (2005) Computer-aided diagnosis of the solitary pulmonary nodule. *Acad Radiol* 12: 570–575.
191. Sidiropoulos KP, Kostopoulos SA, Glotsos DT, Athanasiadis EI, Dimitropoulos ND, et al. (2013) Multimodality GPU-based computer-assisted diagnosis of breast cancer using ultrasound and digital mammography images. *Int J Comput Assist Radiol Surg* 8: 547–560.

192. Siebers S, Zenk J, Bozzato A, Klintworth N, Iro H, et al. (2010) Computer aided diagnosis of parotid gland lesions using ultrasonic multi-feature tissue characterization. *Ultrasound Med Biol* 36: 1525–1534.
193. Sivaramakrishna R, Powell KA, Lieber ML, Chilcote WA, Shekhar R (2002) Texture analysis of lesions in breast ultrasound images. *Comput Med Imaging Graph* 26: 303–307.
194. Skogen K, Ganesan B, Good C, Critchley G, Miles K (2013) Measurements of heterogeneity in gliomas on computed tomography relationship to tumour grade. *J Neurooncol* 111: 213–219.
195. Spector DI, Fischetti AJ, Kovak-McClaran JR (2011) Computed tomographic characteristics of intrapelvic masses in dogs. *Vet Radiol Ultrasound* 52: 71–74.
196. Spira D, Sokler M, Vogel W, Löffler S, Spira SM, et al. (2011) Volume and attenuation computed tomography measurements for interim evaluation of Hodgkin and follicular lymphoma as an additional surrogate parameter for more confident response monitoring: a pilot study. *Cancer Imaging* 11: 155–162.
197. Spira D, Wecker M, Spira SM, Hetzel J, Spengler W, et al. (2013) Does volume perfusion computed tomography enable differentiation of metastatic and non-metastatic mediastinal lymph nodes in lung cancer patients? A feasibility study. *Cancer Imaging* 13: 323–331.
198. Stantz KM, Cao M, Cao N, Liang Y, Miller KD (2011) Monitoring the longitudinal intra-tumor physiological impulse response to VEGFR2 blockade in breast tumors using DCE-CT. *Mol Imaging Biol* 13: 1183–1195.
199. Strzelecki M, Materka A, Drozd J, Krzeminska-Pakula M, Kasprzak JD (2006) Classification and segmentation of intracardiac masses in cardiac tumor echocardiograms. *Comput Med Imaging Graph* 30: 95–107.
200. Su Y, Wang Y, Jiao J, Guo Y (2011) Automatic detection and classification of breast tumors in ultrasonic images using texture and morphological features. *Open Med Inform J* 5: 26–37.
201. Sujana H, Swarnamani S, Suresh S (1996) Application of artificial neural networks for the classification of liver lesions by image texture parameters. *Ultrasound Med Biol* 22: 1177–1181.
202. Sun T, Zhang R, Wang J, Li X, Guo X (2013) Computer-aided diagnosis for early-stage lung cancer based on longitudinal and balanced data. *PLoS One* 8: 1–6.
203. Tan S, Kligerman S, Chen W, Lu M, Kim G, et al. (2013) Spatial-temporal [(1)(8)F]FDG-PET features for predicting pathologic response of esophageal cancer to neoadjuvant chemoradiation therapy. *Int J Radiat Oncol Biol Phys* 85: 1375–1382.
204. Thijssen JM, Verbeek AM, Romijn RL, de Wolf-Rouendaal D, Oosterhuis JA (1992) Echographic differentiation of intraocular melanomas by computer analysis. *Acta Ophthalmol Suppl* 204: 26–34.
205. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, et al. (2011) Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* 52: 369–378.
206. Tozer DJ, Jager HR, Danchaivijitr N, Benton CE, Tofis PS, et al. (2007) Apparent diffusion coefficient histograms may predict low-grade glioma subtype. *NMR Biomed* 20: 49–57.
207. Tsantis S, Cavouras D, Kalatzis I, Piliouras N, Dimitropoulos N, et al. (2005) Development of a support vector machine-based image analysis system for assessing the thyroid nodule malignancy risk on ultrasound. *Ultrasound Med Biol* 31: 1451–1459.
208. Tse DM, Joshi N, Anderson EM, Brady M, Gleeson FV (2012) A computer-aided algorithm to quantitatively predict lymph node status on MRI in rectal cancer. *Br J Radiol* 85: 1272–1278.
209. Tuma J, Novakova B, Schwarzenbach HR, Jungius KP, Hollerweger A, et al. (2011) [Image analysis in the differential diagnosis of renal parenchyma lesions]. *Ultraschall Med* 32: 286–292.
210. Vaidya M, Creach KM, Frye J, Dehdashti F, Bradley JD, et al. (2012) Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer. *Radiother Oncol* 102: 239–245.
211. Virmani J, Kumar V, Kalra N, Khandelwal N (2013) A comparative study of computer-aided classification systems for focal hepatic lesions from B-mode ultrasound. *J Med Eng Technol* 37: 292–306.
212. Virmani J, Kumar V, Kalra N, Khandelwal N (2013) SVM-based characterization of liver ultrasound images using wavelet packet texture descriptors. *J Digit Imaging* 26: 530–543.
213. Wang H, Guo XH, Jia ZW, Li HK, Liang ZG, et al. (2010) Multilevel binomial logistic prediction model for malignant pulmonary nodules based on texture features of CT image. *Eur J Radiol* 74: 124–129.
214. Way TW, Sahiner B, Chan HP, Hadjiiski L, Cascade PN, et al. (2009) Computer-aided diagnosis of pulmonary nodules on CT scans: improvement of classification performance with nodule surface features. *Med Phys* 36: 3086–3098.
215. Willaime JM, Turkheimer FE, Kenny LM, Aboagye EO (2013) Quantification of intra-tumour cell proliferation heterogeneity using imaging descriptors of 18F fluorothymidine-positron emission tomography. *Phys Med Biol* 58: 187–203.
216. Win T, Miles KA, Janes SM, Ganesan B, Shastry M, et al. (2013) Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer. *Clin Cancer Res* 19: 3591–3599.
217. Wu H, Sun T, Wang J, Li X, Wang W, et al. (2013) Combination of radiological and gray level co-occurrence matrix textural features used to distinguish solitary pulmonary nodules by computed tomography. *J Digit Imaging* 26: 797–802.
218. Wu WJ, Lin SW, Moon WK (2012) Combining support vector machine with genetic algorithm to classify ultrasound breast tumor images. *Comput Med Imaging Graph* 36: 627–633.
219. Yan K, Yu Y, Tinney E, Baraldi R, Liao L (2012) Clinical study of a noninvasive multimodal sono-contrast induced spectroscopy system for breast cancer diagnosis. *Med Phys* 39: 1571–1578.
220. Yang X, Mrozek E, Lustberg M, Jia G, Sammet S, et al. (2012) Microcirculatory fraction (MCF(I)) as a potential imaging marker for tumor heterogeneity in breast cancer. *Magn Reson Imaging* 30: 1059–1067.
221. Yoshida H, Casalino DD, Keserci B, Coskun A, Ozturk O, et al. (2003) Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumours in ultrasound images. *Phys Med Biol* 48: 3735–3753.
222. Yoshida H, Nappi J, MacEneaney P, Rubin DT, Dachman AH (2002) Computer-aided diagnosis scheme for detection of polyps at CT colonography. *Radiographics* 22: 963–979.
223. Yoshiura K, Miwa K, Yuasa K, Tokumori K, Kanda S, et al. (1997) Ultrasonographic texture characterization of salivary and neck masses using two-dimensional gray-scale clustering. *Dentomaxillofac Radiol* 26: 332–336.
224. Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, et al. (2009) Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn Reson Med* 62: 1609–1618.
225. Zhang X, Kanematsu M, Fujita H, Zhou X, Hara T, et al. (2009) Application of an artificial neural network to the computer-aided differentiation of focal liver disease in MR imaging. *Radiol Phys Technol* 2: 175–182.
226. Zheng Y, Englander S, Baloch S, Zacharaki EI, Fan Y, et al. (2009) STEP: spatiotemporal enhancement pattern for MR-based breast tumor diagnosis. *Med Phys* 36: 3192–3204.
227. Zhou S, Shi J, Zhu J, Cai Y, Wang R (2013) Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image. *Biomed Signal Process Control* 8: 688–696.
228. Zhu Y, Tan Y, Hua Y, Wang M, Zhang G, et al. (2010) Feature selection and performance evaluation of support vector machine (SVM)-based classifier for differentiating benign and malignant pulmonary nodules by computed tomography. *J Digit Imaging* 23: 51–65.
229. Liu H, Motoda H (2008) Computational methods of feature selection; Liu H, Motoda H, editors. Boca Raton: Chapman & Hall/CRC.