European Journal of Radiology 127 (2020) 108991

Contents lists available at ScienceDirect

European Journal of Radiology

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

A review of original articles published in the emerging field of radiomics

Jiangdian Song (PhD)[a,b,1,*], Yanjie Yin (MS)[a,1], Hairui Wang (Dr)[c], Zhihui Chang (Dr)[c], Zhaoyu Liu (Dr)[c], Lei Cui (Dr)[a]

[a] School of Medical Informatics, China Medical University, No.77 Puhe Road, Shenyang North New Area, Shenyang, 110122, Liaoning Province, PR China
[b] School of Medicine, Department of Radiology, Stanford University, 1201 Welch Rd Lucas Center PS055, Palo Alto, CA 94306, United States
[c] Department of Radiology, Shengjing Hospital of China Medical University, No.36 Sanhao Street, Heping District Shenyang, 110004, Liaoning, Province, PR China

ARTICLE INFO

Keywords:
Radiology
Bibliometrics
Radiomics
Hotspot analysis

ABSTRACT

Purpose: To determine the characteristics of and trends in research in the emerging field of radiomics through bibliometric and hotspot analyses of relevant original articles published between 2013 and 2018.
Methods: We evaluated 553 original articles concerning radiomics, published in a total of 61 peer-reviewed journals between 2013 and 2018. The following information was retrieved for each article: radiological subspecialty, imaging technique(s), machine learning technique(s), sample size, study setting and design, statistical result(s), study purpose, software used for feature calculation, funding declarations, author number, first author's affiliation, study origin, and journal name. Qualitative and quantitative analyses were performed for the manually extracted data for identification and visualization of the trends in radiomics research.
Results: The annual growth rate in the number of published papers was 177.82% ($p < 0.001$). The characteristics and trends of research hotspots in the field of radiomics were clarified and visualized in this study. It was found that the field of radiomics is at a more mature stage for lung, breast, and prostate cancers than for other sites. Radiomics studies primarily focused on radiological characterization (215) and monitoring (182). Logistic regression and LASSO were the two most commonly used techniques for feature selection. Non-clinical researchers without a medical background dominated radiomics studies (70.52%), the vast majority of which only highlighted positive results (97.80%) while downplaying negative findings.
Conclusions: The reporting of quantifiable knowledge about the characteristics and trajectories of radiomics can inform researchers about the gaps in the field of radiomics and guide its future direction.

1. Introduction

The rapidly evolving field of radiomics has attracted considerable interest in the field of radiology- and clinical oncology-related disciplines in the past few years [1–4]. Using extensive medical imaging features extracted from regions of interest (ROIs) on medical images for the purpose of heterogeneity assessments, radiomics has demonstrated its value for improving clinical practice concerning human diseases, particularly cancers [5–7]. For now, radiomics studies allow for the integration of clinical biomarkers, genetic and biochemical indices, and image-based heterogeneity signatures into one single model for establishing the clinical diagnosis and prognosis of human diseases [8,9].

The aim of a radiomics study is generally based on a specific challenge or hypothesis that has existed in current clinical practice. To provide a radiomics-based solution, patients scanned by specific imaging modalities are or will be enrolled in accordance with the study's inclusion criteria. Computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI) are the most commonly used imaging modalities. Next, experts/radiologists perform manual or automatic segmentation of regions of interest (ROIs) on the images. Automatic programs are subsequently designed for high-throughput extraction of quantitative imaging features from the ROIs. Once all the lesion features are extracted, the high-throughput feature data matrix and clinical indicators are integrated into predictive models for the development of solutions specific to the study subject. Finally, the proposed model is internally/externally validated in order to ensure

Abbreviations: AGR, annual growth rate; CT, computed tomography; CAD, computer-aided diagnosis; LASSO, least absolute shrinkage and selection operator; gCLUTO, Graphical Clustering Toolkit; MeSH, Medical Subject Headings; MRI, magnetic resonance imaging; PET, positron emission tomography; ROI, region of interest; US, ultrasonography

* Corresponding authors at: 1201 Welch Rd Lucas Center PS055, Palo Alto, CA 94306, United States.
E-mail addresses: song.jd0910@gmail.com (J. Song), lcui@cmu.edu.cn (L. Cui).

[1] these authors contributed equally to the work.

https://doi.org/10.1016/j.ejrad.2020.108991
Received 6 December 2019; Received in revised form 6 March 2020; Accepted 3 April 2020

accuracy [10–12].

Unlike conventional computer-aided diagnosis (CAD), radiomics focuses on providing information about human diseases from deeper and latent perspectives, and the identification of high-dimensional heterogeneity information of images by radiomics is unmatched by traditional CADs [13]. Through the discovery and accumulation of subtle details that cannot be identified by visual analysis of conventional imaging studies, information potentially conveyed by medical images, such as the severity, progression, or recurrence of a disease, is further interpreted by radiomics studies [14–17]. Of late, radiomics studies involving the prediction of lymphatic metastasis from colorectal cancer [18], evaluation of long-term survival in nasopharyngeal carcinoma [19], proper treatment selection for non-small cell lung cancer [20], and evaluation of gene expressions [21,22] have shown potential clinical value.

However, the current development of radiomics faces challenges. An abundance of radiomic studies has been published in the last few years. However, there has been insufficient oversight in the manner in which they are reported, analyzed, and conducted. Besides, the current status of radiomics research on various clinical topics remains unclear, which makes it difficult to identify the research progress of a specific topic in radiomics. Finally, the need for panoramic and comprehensive statistical analyses of all evidence in each study, with the reporting and analysis of both positive and negative results [23], has been proposed, although it is rarely implemented.

To date, radiomics has been applied to disease prevention, diagnosis, efficacy evaluation, and prognosis prediction. However, quantifiable knowledge about research concerning radiomics in terms of recent developments, current trends, and future trajectories is scarce. Accordingly, the aim of the present study was to determine and present the characteristics of and trends in research in the emerging field of radiomics through bibliometric and hotspot analyses of relevant original articles published in this field.

2. Materials and Methods

This was a retrospective study that did not involve human subjects; therefore, the need for institutional review board approval was waived.

We searched PubMed for original radiomics studies published between January 1st, 2012 and December 31st, 2018 using “radiomics” OR “radiomic” as the search term. Abstracts, case reports, review articles, clinical perspectives, state of the art studies, editorials, letters, technical notes, quizzes, audiovisual media, educational material, book reviews, commentaries, and news were excluded (the detailed search strategy is shown in Supplementary Figure S1). Original articles were defined as studies that involved radiomics-related objectives or hypotheses and contained specifically articulated methods and results sections. All included articles were obtained from the journal websites. Both print and online-only original articles registered in online archives were included.

We first performed a bibliometric analysis of all included articles in order to retrieve the features of studies in this field. The classification of radiological subspecialties for the purpose of bibliometric analysis has been well defined in previous studies [24,25]. For analysis of the characteristics of and trends in radiomics studies, the following information was extracted from each article: radiological subspecialty [neuroradiology, head and neck, thyroid, breast, thoracic, cardiac, abdominal, musculoskeletal, genitourinary, or miscellaneous (not conforming to one of the aforementioned categories)], imaging technique (s) [US, CT, MRI, PET, mixed (more than one radiological technique), or other], sample size (≤50, 51–100, or > 100), study setting (single center or validated on multiple centers), study design (prospective or retrospective, training/validation schemes), statistical analysis (used or not used, positive or negative result), study purpose [image acquisition, pre-processing, detection, characterization, monitoring (prediction of therapy response, risk assessment, and prognosis), or reporting, according to consensus in the radiology field [26]], funding sources (government, private, or other), author number (< 4, 4–7, > 7), first author’s affiliation [radiology (including radiology, nuclear medicine, and other imaging-related specialties), medicine or related specialties (including internal medicine, pediatrics, psychiatry, neurology, dermatology, etc.), surgery or related specialties (including surgery, obstetrics and gynecology, orthopedics, anesthesiology, pathology, etc.), or other (including basic science, laboratory, or other research institutes)], software or toolkits used for radiomics feature calculation, study origin (for the purpose of our research, the country in which the first author’s institution was located was considered the study origin [25]), journal name, and specific machine learning techniques used for feature selection.

We then performed hotspot analysis to identify research topics in the radiomics field. All the main Medical Subject Heading (MeSH) terms/subheadings for each article were extracted by the Bibliographic Item Co-Occurrence Matrix Builder software, following which a co-occurrence matrix was constructed to clarify the co-occurrence of these terms in the articles [27]. Subsequently, the Graphical Clustering Toolkit (gCLUTO) software [28] was used to perform visualized clustering of the main MeSH terms/subheadings, and a Mountain Visualization map and visual co-occurrence matrix were created.

A strategy diagram was constructed on the basis of the clustering result to understand the state of each research topic and the relationship between the various topics in this field. The x-axis of the strategy diagram represented the centrality of the cluster, i.e. whether the cluster was highly correlated with other clusters (the right side indicated high centrality, which showed that the cluster had higher intercluster associations with other clusters). The y-axis represented the density, i.e., the intra-cluster correlation of each cluster, which referred to the maturity of cluster development (the upper side indicated higher intra-cluster correlation of the cluster). Finally, the keywords mentioned in all the articles were used for further co-occurrence cluster analysis using CiteSpace [29]. This allowed us to obtain data regarding historical changes in research topics.

Seven investigators, including three radiologists with > 10 years of experience in radiology and four bibliometric experts working in the Department of Medical Informatics, independently reviewed all the articles. First, data were manually collected by the bibliometric experts. Then, the radiologists performed a blind review of the data. All information was manually entered into a local Excel spreadsheet, and the final internal quality check was performed by Dr. Jiangdian Song and (Yanjie Yin). Any disagreements were resolved in a consensus meeting with the radiologists.

Variance analysis was used to examine differences and trends in articles categorized by the radiological subspecialty, imaging technique (s), sample size, and other variables. A p-value was derived to determine the statistical significance of trends ($p < 0.05$ was considered statistically significant). All statistical analyses were performed using R software, version 3.4.3.

3. Results

In total, 690 articles were retrieved by our search strategy. Of these, 553 original articles were considered eligible, and 137 articles that did not meet the inclusion criteria were excluded (Supplementary Figure S1). The annual growth rate (AGR) of relevant original articles published from 2013 to 2018 was 177.82% ($p < 0.001$, ANOVA; Supplementary Figure S2). The included articles were published in a total of 61 peer-reviewed journals, of which *European Radiology* (n = 50) and *Scientific Reports* (n = 49) were the most common.

The majority of the studies were single-center studies (n = 445, 80.47%). The median number of participants in the studies was 85, and only 48.10% studies had a sample size greater than 100. However, the average sample size increased to 200 in 2018. Detailed statistics for the sample size with the training/validation schemes are presented in

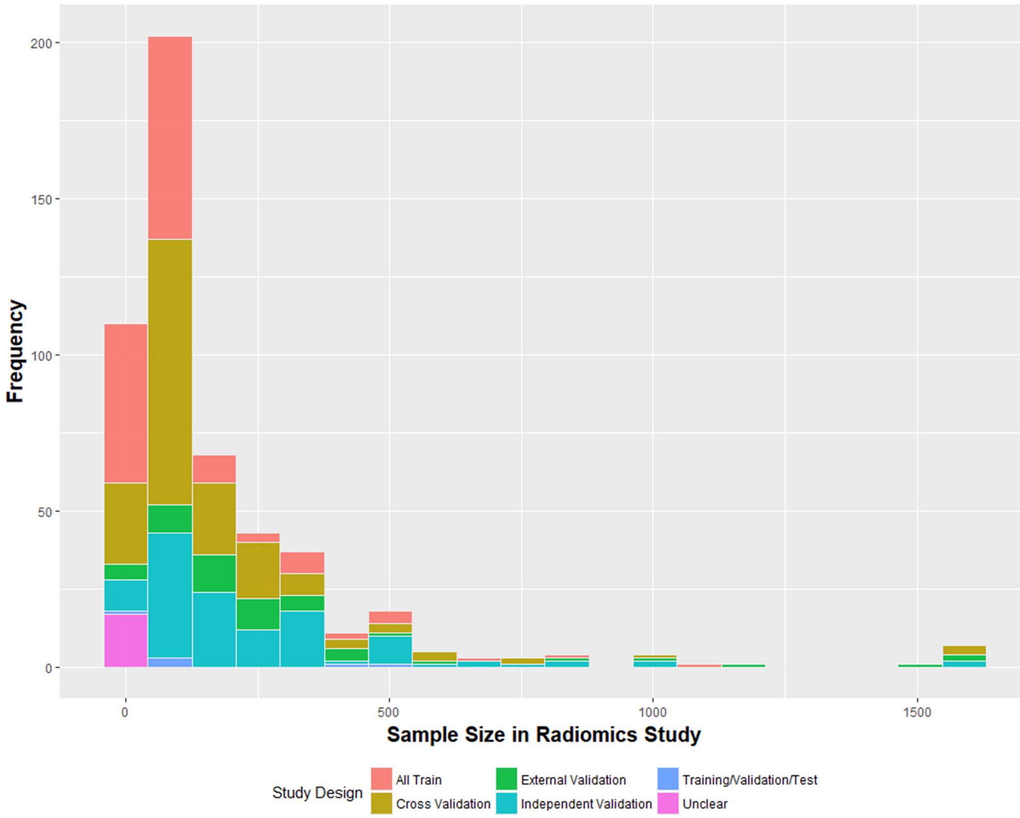

Fig. 1. Sample size statistics with training/validation schemes in studies published in the field of radiomics (2013–2018). Five studies reported that more than 2,000 CT images were used, although they did not report the actual number of included patients. Therefore, we removed these abnormal points to ensure that all other data columns were clearly displayed in the histogram.

Fig. 1. Overall, 60.76% studies included more than seven authors. Statistical analysis was described in 90.24% studies, and 97.80% studies reported positive radiomics results. The first author was a radiology expert in 163 studies (29.48%), and 387 studies (68.98%) were supported by funding from the government or other entities, with a disparity in the number of studies that received funding between the United States and other countries. While 80.87% studies conducted in the United States were funded, only 50% studies conducted in other countries received funds. Another result from this study is closely related to the statistical numbers: most radiomics studies originated in the United States (32.91%), with an AGR of 158.28%, with other common study origins including China (29.84%; AGR, 387.70%), Italy (n = 27), and Canada (n = 27).

We found that CT was the most frequently used imaging modality (43.56%), followed by MRI (35.93%). The increasingly frequent use of MRI for radiomics was an interesting finding, with the AGR for studies using MRI being the highest (196.25%) among those studies concerning other modalities (Table 1). In addition, the proportion of studies with a sample size of > 100 gradually increased with time (Table 1); 32.02%, 36.72%, 42.36% and 53.78% studies had sample sizes greater than 100 in 2015, 2016, 2017, and 2018, respectively. Studies with multi-center validation(s) were associated with an AGR of 114.06%, whereas prospective studies were associated with an AGR of 169.81%.

Table 1
Radiology subspecialties (major reported subspecialty) evaluated in original articles published in the field of radiomics (2013–2018)

Radiological subspecialty	2013	2014	2015	2016	2017	2018	Total
Neuroradiology	/	/	1	8	28	72	109
Head and neck	1	/	3	1	17	53	75
Thyroid	/	/	/	/	1	/	1
Breast	/	/	4	4	16	23	47
Thoracic	1	2	10	21	40	73	147
Cardiac	/	/	/	/	1	2	3
Abdominal	/	/	3	7	12	48	70
Musculoskeletal	/	/	/	/	4	2	6
Genitourinary	/	/	3	3	14	31	51
Miscellaneous	/	/	1	5	11	27	44
Total	2	2	25	49	144	331	553

The numbers of original articles increased in various radiologic subspecialties: 26.71% of the published articles were on thoracic diseases, the most widely studied topic. The fastest growing subspecialty was neuroradiology, with an AGR of 316.02% for published articles. In addition to common human diseases, studies in the miscellaneous category were also evaluated (n = 44, 7.94%). These included studies on Meniere's disease, lymph node disease, melanoma, xerostomia, and other diseases. Besides human diseases, radiomics studies also included feature descriptions, tumor heterogeneity assessments, image filtering, and model evaluation. Table 2 presents the details of published original articles classified by the subspecialty.

Regarding the radiological areas of interest in radiomics studies, original articles involving characterization and monitoring of lesions were the most common (n = 214 and 183, respectively), followed by those involving detection and pre-processing (n = 99 and 38, respectively). Studies involving image acquisition and reporting were rare. With respect to machine learning techniques, logistic regression and least absolute shrinkage and selection operator (LASSO) were the most commonly used techniques for feature selection (n = 69 and 69, respectively), followed by random forest, Cox regression, coefficient, and support vector machine (n = 51, 49, 41, and 38, respectively).

With regard to the statistical software and toolkits used for feature calculation, more than half of the studies reported detailed feature extraction methods (302 studies), whereas 251 studies stated that an in-house software was used or did not report the software. Matlab was the most frequently used tool for radiomic feature calculation (159 studies), followed by open access feature extraction toolkits such as Pyradiomics (29 studies) and Imaging Biomarker Explorer (23 studies; Fig. 2).

By using hotspot analysis, a total of 504 main MeSH terms/subheadings were extracted. An optimal threshold of six was used to intercept high-frequency terms, and a total of 36 main MeSH terms/

Table 2
Features of original articles published in the field of radiomics (2013–2018)

Feature	2013	2014	2015	2016	2017	2018	Total
Study purpose							
Image acquisition	/	/	1	/	/	1	2
Pre-processing	1	1	1	5	9	21	38
Detection	1	/	2	9	29	58	99
Characterization	/	/	11	22	43	138	214
Monitoring	/	1	10	13	63	96	183
Reporting	/	/	/	/	/	/	0
Imaging technique							
CT	1	2	15	25	55	142	240
MRI	/	/	5	15	48	130	198
PET	1	/	3	3	9	48	64
Sample size							
≤50	2	1	11	22	46	89	171
51–100	/	/	6	9	37	64	116
> 100	/	1	8	18	61	178	266
Machine learning technique							
Logistic regression	/	/	1	3	16	49	69
LASSO	/	/	1	4	11	53	69
random forest	/	/	2	3	10	36	51
Cox regression	/	1	1	4	15	28	49
coefficient	1	/	4	10	9	17	41
SVM	/	/	2	4	5	27	38
PCA	/	/	/	2	8	14	24
linear	/	/	2	2	2	7	13
ANOVA	/	/	/	/	2	4	6
Bayes	/	/	/	/	2	2	4
Study design							
All train	2	1	9	25	53	55	145
Cross validation	/	/	9	11	47	109	176
Independent validation	/	/	2	8	26	91	127
External validation	/	1	2	3	4	45	55
Training/validation/test	/	/	/	/	3	3	6
Unclear	/	/	3	2	11	28	44

Only the top ranked items are listed. SVM: support vector machine. PCA: principal component analysis. LASSO: least absolute shrinkage and selection operator. Coefficient item represents the studies only declare the coefficient-based method in the text. Since this group is complicated, hence we summarize it with coefficient (including coefficient of variation, intra-class correlation coefficients, correlation coefficients, concordance correlation coefficients, overall coefficient of variation, and Spearman correlation coefficient, *et, al.*). Bayes: Bayesian neural network. Linear item represents the studies declare that the linear discriminant analysis is used. Independent validation denotes validation dataset from the same institution, and the external validation denotes validation dataset from different institutions.

subheadings were identified (accounting for 41.48% of all MeSH terms/subheadings in the included papers), as shown in Fig. 3. Eventually, five clusters were created, as shown by the Mountain Visualization in Fig. 3B. Appearance of terms within the same cluster indicated that those terms co-occurred more frequently, and the research topics of each cluster could be determined according to the terms in this cluster. Cluster 0 included studies related to neuroradiology. Cluster 1 included studies related to lung neoplasms, with a general focus on pathology, mutations, and radiotherapy. Cluster 2 included studies evaluating breast and prostate cancers using machine learning techniques. Cluster 3 included studies related to head and neck cancer, squamous cell cancer, and PET or PET/CT images. Finally, Cluster 4 included studies emphasizing engineering approaches, particularly for lung cancer diagnosis. The cluster mountain map and visual co-occurrence matrix are presented in Fig. 3, which indicates that the intra-cluster correlation was the highest for clusters 0 and 1, which exhibited the lowest standard deviation values.

A strategic diagram based on the clustering results indicated that the topics in clusters 1 and 2 not only showed high intra-cluster correlation but also high inter-cluster correlation, as shown in Fig. 4. Although Cluster 0 exhibited good intra-cluster correlation, it was not closely correlated with other clusters and was consequently not at the core of radiomics. We also found that the studies in Cluster 4 had received widespread attention and were closely connected with other clusters, but the correlation of the subtopics within this cluster was poor. The topics in Cluster 3 showed neither high intra-cluster correlation nor close inter-cluster correlation.

For further visualization of trends in research topics, all the keywords of the enrolled articles were used in keyword co-occurrence analysis. The results are presented in Fig. 5. Prevailing research topics in this field changed on an annual basis. The topic of radiogenomics was highlighted in 2015, whereas machine learning methods based on CT images of lung cancer, prostate cancer, and glioblastoma were mainstream in 2016. In 2017, studies focused on texture analysis using MRI, primarily in relation to survival predictions for patients with breast and liver cancers. In 2018, the subjects tended to diversify, with studies about CT images, survival prognosis, and gliomas being more prominent.

We also prepared a relationship map via a dynamic network to illustrate all co-operative relationships among the authors. In this map, interactions are available if the reader wants to get more information from the dynamic network. For display clarity, a threshold (co-operation count of > 3) was used in the network. Detailed information regarding this network is provided at: https://mi-12.github.io/Radiomics-Author-Network.github.io/ (open access).

4. Discussion

In the present study, we found that the field of radiomics has been undergoing rapid development over the past 5 years, with an AGR of 177.82% for original published articles. We also identified the common characteristic and the predominant topics of studies in the emerging field of radiomics over the last few years. It was also found that radiomics research of lung, breast, and prostate cancers is more developed than research in other areas. The two most common study objectives involved the characterization and monitoring of disease, and logistic regression and LASSO were the most common techniques used to address the relevant clinical problems. With the gradual standardization of study protocol and design [6], the number of studies with multi-center validation and prospective studies has increased significantly. This indicates that future investigators should yield radiomics studies with a higher quality of evidence.

In the present study, original articles with more than seven authors accounted for the majority of radiomics studies (n = 336, 60.76%), probably because the radiomics workflow (patient registration, standardization and pre-processing of patient data, ROI segmentation, feature extraction and model construction, and external validation) require adequate expertise and considerable effort to ensure the accuracy of the study results. However, because the qualifications of the staff members involved in radiomics studies are uncertain, we recommend that investigators involved in future radiomics studies the appropriate and extensive training in the field.

The clustering results of our study indicated that radiomics studies related to lung, breast and prostate cancers showed more tense inter-class correlation and intra-class correlation and were more developed than studies on neuroradiology and head and neck and squamous cell cancers. The primary reason for the loose correlation between studies related to neuroradiology and other studies is that brain cancer is not generally analyzed in combination with other cancers, with the primary focus being on primary brain tumors such as glioblastoma [8,30]. Thus, the subtopics within Cluster 0 are highly correlated. However, studies have confirmed that it is possible to analyze lung cancer, for example, together with other cancers [4]. The topic of engineering methods is within Cluster 4 in the strategic diagram, which means that it is highly correlated with other clusters. This is quite plausible since almost all radiomics studies use engineering analysis methods. However, studies on engineering method research are rare in the field of radiomics, which results in its poor intra-class correlation. Studies related to

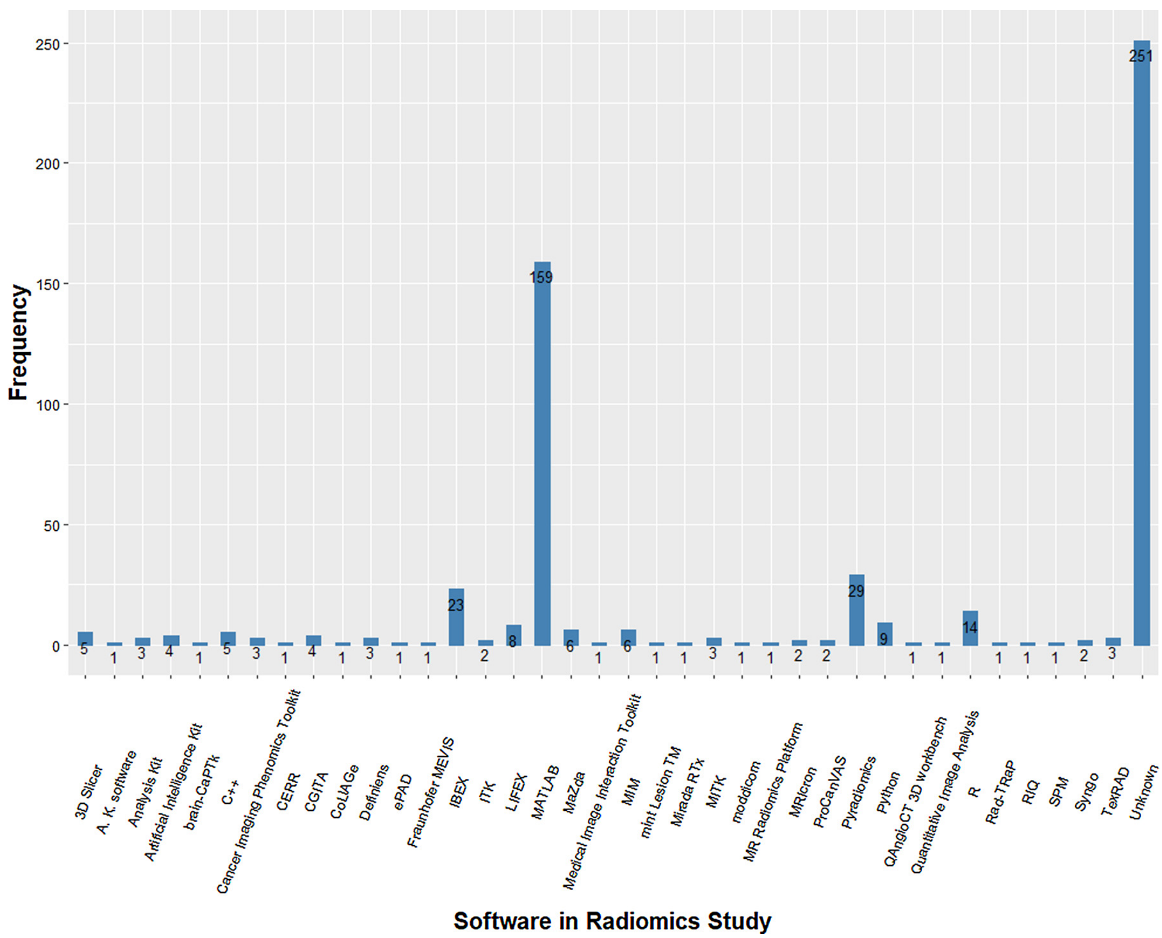

Fig. 2. Statistics regarding the software and toolkits used for feature calculation in studies published in the field of radiomics (2013–2018). Nine studies used more than one feature extraction software while 251 reported the use of in-house feature calculation methods (marked as unknown). CERR: Computational Environment for Radiological Research, CGITA: Chang-Gung Image Texture Analysis, CoLIAGe: Co-occurrence of Local Anisotropic Gradient Orientations, IBEX: Imaging Biomarker Explorer, MIM: MIMMaestro workstation, RIQ: CERR Radiomics Imaging Quantification toolbox.

squamous cell and head and neck cancers and PET/CT images showed low inter-class and intra-class correlations, suggesting that researchers in these subtopics should pay attention to the incorporation and fusion of methods and topics from other fields in the future.

Logistic regression and LASSO are the top two machine learning methods in radiomics. This is consistent with the finding that characterization and monitoring are the top two most frequently covered oncology topics in radiomics (n = 214 and 183, respectively). Because radiomics often involves screening of hundreds of feature vectors, logistic regression is the primary tool used for classification in characterization, while LASSO is the most useful tool for feature selection in both classification and survival prognosis determination.

We found that approximately 70% studies on radiomics had received funding; this rate was higher than that (23.0%) observed for interventional/radiology studies published in major American journals [31]. The United States leads the world in the number of medical research publications[25]; this also holds true for radiomics studies. Our results also indicated that researchers from 23 countries conducted studies on radiomics, however, a comparison of funding rates between the United States and other countries clearly showed that American studies received more funding. Accordingly, we infer that government funding is an important factor for promoting radiomics research.

Our review showed that many radiomic studies explored gene expression, a topic that is key to the field of genomics [32,33]. It has been recently proposed that the scope of radiomics should be expanded by considering imaging, genomics, and clinical biochemistry, since the integration of multidisciplinary information could add more clinical value [34]. One study proposed a certain correlation between radiomic phenotypic features and gene expression in non-small cell lung cancer [4]. Another study suggested that the ability of radiomic features to predict response varied across different receptor subtypes [35]. Recently, a study indicated that the correlation between epidermal growth factor receptor expression and tumor invasion can be demonstrated by radiomics methods [36]. The increasing innovation in genetic and radiomics technologies suggests that future researchers should pay more attention to this interdisciplinary research field. The integration of radiomics, radiogenomics, and clinical biomarkers is a promising stratagem for more precise auxiliary diagnosis and treatment of human diseases in the future [37–39].

Traditional texture analysis has long been applied for image analysis; however, the modern term of radiomics was first proposed in the context of medical imaging feature analysis in 2012 [11,40,41]. Studies related to radiomics that were published in 2012 were not original articles, therefore, the literature search was initiated from 2013. Current radiomics study is mainly aimed at the clinical problems of human beings, and proposes solutions from the perspective of medical imaging. Therefore, we only searched the PubMed database for retrieving studies. Although topics related to engineering methods such as machine learning were highlighted in 2016, we observed a transition to assessments in clinical oncology in the following 2 years. This was a significant trend, suggesting that radiomics studies focus on not only tools but also clinical practice, and that higher clinical evidence levels are increasingly required. Moreover, from 2013 to 2018, both multi-center and prospective studies exhibited a high AGR. Thus, future studies

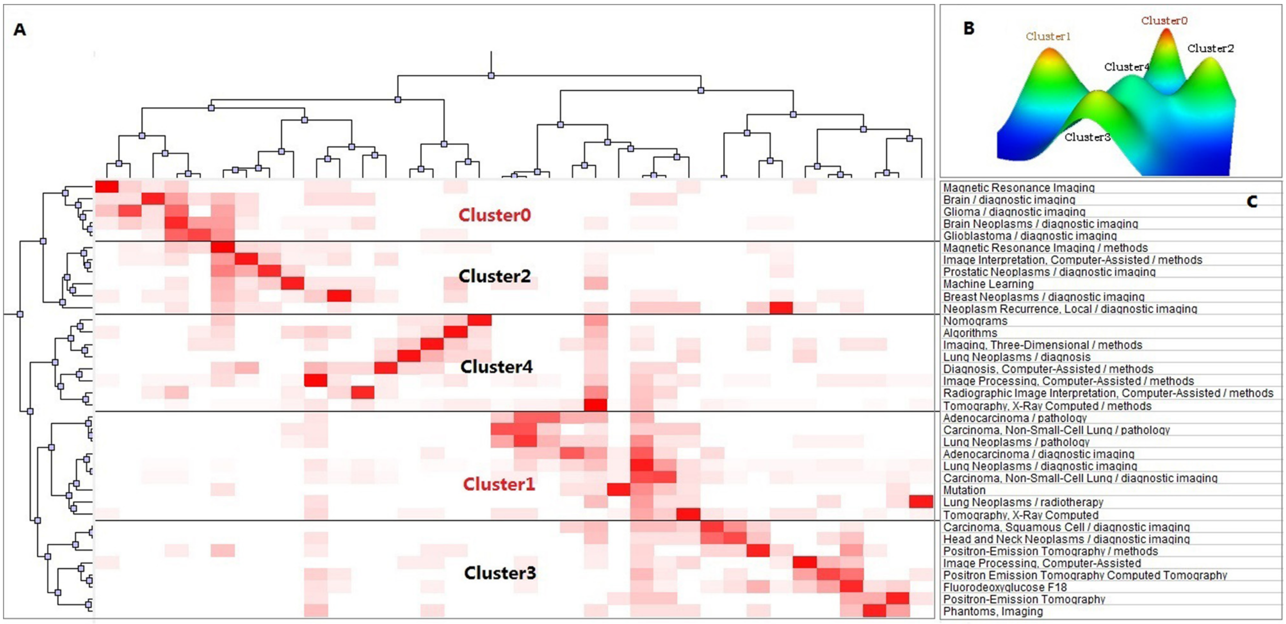

Fig. 3. The visual co-occurrence matrix (A), cluster mountain map (B), and 36 high-frequency main Medical Subject Headings (MeSH) terms/subheadings (C) used in the present review of original articles published in the emerging field of radiomics. Both of the x-axis and y-axis in the co-occurrence matrix (A) represent the 36 high-frequency main MeSH terms/subheadings (the row labels from left to right are the same as the column labels from the top to the bottom), while the red block represents the co-occurrence between two corresponding words. The height of each cluster in (B) represents the internal similarity of the cluster. The volume of each cluster represents the number of included MeSH terms/subheadings. The colour of the cluster peak denotes the standard deviation within the cluster; red indicates a small deviation, while blue indicates a large deviation.

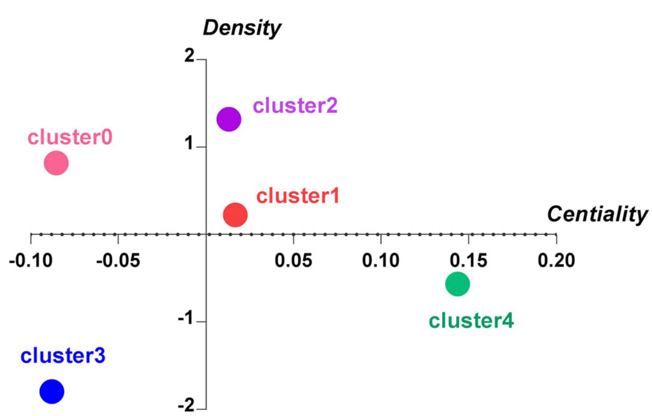

Fig. 4. Strategic diagram of the original articles published in the field of radiomics (2013–2018). Clustering 1 and 2 in the first quadrant indicated that the intra-cluster correlation of the topics in these clusters is high, which denotes the topics are well developed. And their inter-cluster correlation with other clusters are also high, and hence are at the core of the field of radiomics. Clustering 0 in the second quadrant indicates that the topics are not closely related to other clusters. Clustering 3 in the third quadrant indicates that neither the inter-class correlation nor the intra-class correlation of the topics is good. Clustering 4 in the fourth quadrant indicates that although the topics are closely related to other clusters, the intra-class correlation of the topics is however loose.

should perform radiomics experiments with a much higher quality of evidence in order to ensure clinical applicability.

Although there is great momentum in developments in radiomics, potential concerns exist. One of the most serious problems is the reliability of the abundant imaging-derived signatures proposed by radiomics [42]. Reproducibility of radiomics has been demonstrated to be vital for its use in clinical practice and many studies have put effort into testing the reproducibility of radiomics results [43–45]. However, in this study, we found that there is currently a lack of authentication on the reliability of radiomics signatures in radiomics research (single-center studies accounted for the majority of studies). The verification of radiomics signature reproducibility is a factor that cannot be ignored in order to achieve clinical utility of radiomics. Additionally, studies highlight evidence that corroborates with the information from the proposed radiomic signatures while detracting from evidence that contradicts them [23]. Currently, radiomics is in the incipient research stage, and there are few examples where the results of radiomic studies have achieved practical clinical utility. However, the clinical significance of radiomics studies should be considered cautiously with consideration of false-negative results from its techniques [23,46], in order to transform it into a tool for clinical practice in the future.

Another potential concern is the background of many of the researchers in radiomics. Only 29.48% of the original articles included in this study had first authors who were radiology experts. This is in contrast to a previous study that found that 87.6% of the first authors of radiology studies were experts in radiology or related fields [25]. This is because radiomics studies generally focus on one specific clinical issue, particularly cancers, which require the integration of knowledge from radiologists, oncologists, programming engineers, and experts from other fields. Although the involvement researchers outside the field of medicine enriches this area of study by providing alternative views, it can potentially lead to poorly designed studies due to a lack of necessary medical knowledge. Therefore, future radiomics studies should be carefully designed and monitored by people with a medical background to ensure they are carried out in strict accordance with rigorous radiology standards.

Our findings provide suggestions for a more reliable radiomics practice in the future. First, attention should be paid to improving the quality of evidence in future radiomics studies. Increased sample sizes are not only a current trend, but also important for achieving more reliable findings in radiomics studies. According to the sample size data, a minimum sample size of 200 cases was considered appropriate for the present study. Independent or external validation is vital to ensure a high level of evidence. The results of our study indicated that studies with independent/external validation dataset(s) increased to 41% in 2018. Besides, appropriate and extensive training for the staff

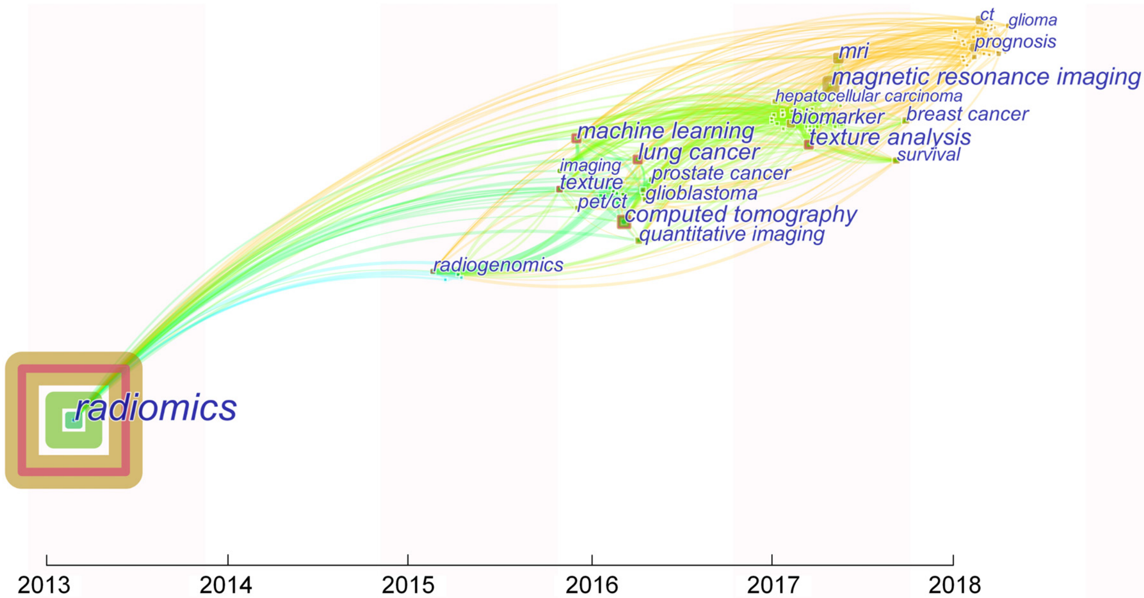

Fig. 5. Timezones based on the keywords in 553 original articles published in the field of radiomics (2013-2018). The CiteSpace software was used to analyse the keywords in studies from each year in order to create a keyword co-occurrence matrix for timezone generation. The larger is the node area, the higher is the keyword frequency. The wider is the outer circle of the node, the greater is the keyword importance. The blue, green, fluorescent green, and yellow lines represent the studies in 2015, 2016, 2017, and 2018, respectively. Since the number of studies in this field has increased dramatically since 2017, research hotspots have also diversified, and the area of each node has gradually become smaller.

members involved in radiomics studies is essential because majority of investigators are from non-medical disciplines and their qualifications are uncertain. Another worthwhile practice is the use of open access feature extraction software such as Pyradiomics and Imaging Biomarker Explorer, which have proven useful for improving the repeatability of imaging features [17,47]. The detailed feature extraction scheme should be published if open access software is not available. In addition, according to the results of the strategic diagram in this study, the inter-class and intra-class correlations for subtopics related to different human parts should be further considered in future radiomics studies. For example, the results of the study in the third quadrant of the strategic diagram indicate that attention should be paid to the integration and fusion of methods and subtopics from other fields in the future. Finally, more rigorous validation of radiomics features, including both positive and negative results, should be implemented for the reproducibility of these features.

This study has some limitations. All data were manually retrieved, and despite independent blind reviews, bias caused by subjective experience was inevitable. The institutional affiliation of the first author was considered the study origin in this study. Although previous studies have shown that the first author is the most meaningful contributor, the contribution of all authors should be considered [48]. Additionally, our search was restricted to the PubMed database. Although the majority of radiomics research studies can be retrieved from the PubMed database, some papers published in journals based outside of the United States and Europe may not be present in this database and additional databases should be used in future research. Finally, we did not document the imaging features extracted in every included study. Further studies should determine the role of specific imaging features in the clinical evaluation of human cancers.

In summary, we conducted a systematic literature review to report the trends in the characteristics and content of studies published in the field of radiomics. The diversified development of radiomics suggests the need for future researchers to provide higher levels of clinical evidence to solve clinical problems as well as the need for a more rigorous experimental design and attention to negative results. Knowledge about the characteristics and future trajectories of radiomics will drive the development of radiology, which promises to provide more valuable information for the diagnosis and treatment of human diseases.

Declaration of Competing Interest

None

Funding

This work was supported by grants China Postdoctoral Fund (2018M630310) and China Scholarship Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgement

We thank all the investigators (Dr. Qinlai Huang, Dr. Baihui Yan, Dr. Mei Li, Dr. Yingxin Xu, and Dr. Yanjie Yin) for their help and contribution to the study. This work was supported by grants China Postdoctoral Fund (2018M630310) and China Scholarship Council. They have made outstanding contributions to the literature retrieval and information extraction. We would like to thank Editage (www.editage.cn) for English language editing.

References

[1] K. Wang, X. Lu, H. Zhou, Y. Gao, J. Zheng, M. Tong, C. Wu, C. Liu, L. Huang, T. Jiang, F. Meng, Y. Lu, H. Ai, X.Y. Xie, L.P. Yin, P. Liang, J. Tian, R. Zheng, Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study, Gut 68 (4) (2019) 729–741.

[2] Z. Liu, X.Y. Zhang, Y.J. Shi, L. Wang, H.T. Zhu, Z. Tang, S. Wang, X.T. Li, J. Tian, Y.S. Sun, Radiomics Analysis for Evaluation of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer, Clinical cancer research : an official journal of the American Association for Cancer Research 23 (23) (2017) 7253–7262.

[3] Y. Huang, Z. Liu, L. He, X. Chen, D. Pan, Z. Ma, C. Liang, J. Tian, C. Liang, Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer, Radiology 281 (3) (2016) 947–957.

[4] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers,

M.M. Rietbergen, C.R. Leemans, A. Dekker, J. Quackenbush, R.J. Gillies, P. Lambin, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nature communications 5 (2014) 4006.
[5] H.J. Aerts, The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review, JAMA oncology 2 (12) (2016) 1636–1642.
[6] P. Lambin, R.T.H. Leijenaar, T.M. Deist, J. Peerlings, E.E.C. de Jong, J. van Timmeren, S. Sanduleanu, R. Larue, A.J.G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F.M. Mottaghy, J.E. Wildberger, S. Walsh, Radiomics: the bridge between medical imaging and personalized medicine, Nature reviews, Clinical oncology 14 (12) (2017) 749–762.
[7] M. Gerlinger, A.J. Rowan, S. Horswell, M. Math, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, P. Tarpey, I. Varela, B. Phillimore, S. Begum, N.Q. McDonald, A. Butler, D. Jones, K. Raine, C. Latimer, C.R. Santos, M. Nohadani, A.C. Eklund, B. Spencer-Dene, G. Clark, L. Pickering, G. Stamp, M. Gore, Z. Szallasi, J. Downward, P.A. Futreal, C. Swanton, Intratumor heterogeneity and branched evolution revealed by multiregion sequencing, The New England journal of medicine 366 (10) (2012) 883–892.
[8] H. Itakura, A.S. Achrol, L.A. Mitchell, J.J. Loya, T. Liu, E.M. Westbroek, A.H. Feroze, S. Rodriguez, S. Echegaray, T.D. Azad, K.W. Yeom, S. Napel, D.L. Rubin, S.D. Chang, G.R.t. Harsh, O. Gevaert, Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities, Science translational medicine 7 (303) (2015) 303ra138.
[9] O. Gevaert, L.A. Mitchell, A.S. Achrol, J. Xu, S. Echegaray, G.K. Steinberg, S.H. Cheshier, S. Napel, G. Zaharchuk, S.K. Plevritis, Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features, Radiology 273 (1) (2014) 168–174.
[10] H. Aerts, Data Science in Radiology: A Path Forward, Clinical cancer research : an official journal of the American Association for Cancer Research 24 (3) (2018) 532–534.
[11] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G. van Stiphout, P. Granton, C.M. Zegers, R. Gillies, R. Boellard, A. Dekker, H.J. Aerts, Radiomics: extracting more information from medical images using advanced feature analysis, European journal of cancer 48 (4) (2012) 441–446.
[12] V. Kumar, Y. Gu, S. Basu, A. Berglund, S.A. Eschrich, M.B. Schabath, K. Forster, H.J. Aerts, A. Dekker, D. Fenstermacher, D.B. Goldgof, L.O. Hall, P. Lambin, Y. Balagurunathan, R.A. Gatenby, R.J. Gillies, Radiomics: the process and the challenges, Magnetic resonance imaging 30 (9) (2012) 1234–1248.
[13] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: Images Are More than Pictures, They Are Data, Radiology 278 (2) (2016) 563–577.
[14] Q. Li, J. Kim, Y. Balagurunathan, Y. Liu, K. Latifi, O. Stringfield, A. Garcia, E.G. Moros, T.J. Dilling, M.B. Schabath, Z. Ye, R.J. Gillies, Imaging features from pretreatment CT scans are associated with clinical outcomes in nonsmall-cell lung cancer patients treated with stereotactic body radiotherapy, Medical physics 44 (8) (2017) 4341–4349.
[15] H. Li, Y. Zhu, E.S. Burnside, K. Drukker, K.A. Hoadley, C. Fan, S.D. Conzen, G.J. Whitman, E.J. Sutton, J.M. Net, M. Ganott, E. Huang, E.A. Morris, C.M. Perou, Y. Ji, M.L. Giger, MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays, Radiology 281 (2) (2016) 382–391.
[16] J. Song, J. Shi, D. Dong, M. Fang, W. Zhong, K. Wang, N. Wu, Y. Huang, Z. Liu, Y. Cheng, Y. Gan, Y. Zhou, P. Zhou, B. Chen, C. Liang, Z. Liu, W. Li, J. Tian, A New Approach to Predict Progression-free Survival in Stage IV EGFR-mutant NSCLC Patients with EGFR-TKI Therapy, Clinical cancer research : an official journal of the American Association for Cancer Research 24 (15) (2018) 3583–3592.
[17] J.J.M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R.G.H. Beets-Tan, J.C. Fillion-Robin, S. Pieper, H. Aerts, Computational Radiomics System to Decode the Radiographic Phenotype, Cancer research 77 (21) (2017) e104–e107.
[18] Y.Q. Huang, C.H. Liang, L. He, J. Tian, C.S. Liang, X. Chen, Z.L. Ma, Z.Y. Liu, Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer, Journal of clinical oncology : official journal of the American Society of Clinical Oncology 34 (18) (2016) 2157–2164.
[19] B. Zhang, J. Tian, D. Dong, D. Gu, Y. Dong, L. Zhang, Z. Lian, J. Liu, X. Luo, S. Pei, X. Mo, W. Huang, F. Ouyang, B. Guo, L. Liang, W. Chen, C. Liang, S. Zhang, Radiomics Features of Multiparametric MRI as Novel Prognostic Factors in Advanced Nasopharyngeal Carcinoma, Clinical cancer research : an official journal of the American Association for Cancer Research 23 (15) (2017) 4259–4269.
[20] J. Song, Z. Liu, W. Zhong, Y. Huang, Z. Ma, D. Dong, C. Liang, J. Tian, Non-small cell lung cancer: quantitative phenotypic analysis of CT images as a potential marker of prognosis, Scientific reports 6 (2016) 38282.
[21] M. Zhou, A. Leung, S. Echegaray, A. Gentles, J.B. Shrager, K.C. Jensen, G.J. Berry, S.K. Plevritis, D.L. Rubin, S. Napel, O. Gevaert, Non-Small Cell Lung Cancer Radiogenomics Map Identifies Relationships between Molecular and Imaging Phenotypes with Prognostic Implications, Radiology 286 (1) (2018) 307–315.
[22] O. Gevaert, J. Xu, C.D. Hoang, A.N. Leung, Y. Xu, A. Quon, D.L. Rubin, S. Napel, S.K. Plevritis, Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data–methods and preliminary results, Radiology 264 (2) (2012) 387–396.
[23] I. Buvat, F. Orlhac, The dark side of radiomics: on the paramount importance of publishing negative results, J Nucl Med (2019).
[24] L.J. Zhang, Y.F. Wang, Z.L. Yang, U.J. Schoepf, J. Xu, G.M. Lu, E. Li, Radiology research in mainland China in the past 10 years: a survey of original articles published in Radiology and European Radiology, European radiology 27 (10) (2017) 4379–4382.
[25] K.J. Lim, D.Y. Yoon, E.J. Yun, Y.L. Seo, S. Baek, D.H. Gu, S.J. Yoon, A. Han, Y.J. Ku, S.S. Kim, Characteristics and trends of radiology research: a survey of original articles published in AJR and Radiology between 2001 and 2010, Radiology 264 (3) (2012) 796–802.
[26] A. Hosny, C. Parmar, J. Quackenbush, L.H. Schwartz, H. Aerts, Artificial intelligence in radiology, Nature reviews, Cancer 18 (8) (2018) 500–510.
[27] B. Shi, W. Wei, X. Qin, F. Zhao, Y. Duan, W. Sun, D. Li, Y. Cao, Mapping theme trends and knowledge structure on adipose-derived stem cells: a bibliometric analysis from 2003 to 2017, Regenerative medicine 14 (1) (2019) 33–48.
[28] G.K. Matt Rasmussen, gCLUTO: An Interactive Clustering, Visualization, and Analysis System, (2004).
[29] C. Chen, Z. Hu, S. Liu, H. Tseng, Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace, Expert opinion on biological therapy 12 (5) (2012) 593–608.
[30] O. Gevaert, L.A. Mitchell, A.S. Achrol, J. Xu, S. Echegaray, G.K. Steinberg, S.H. Cheshier, S. Napel, G. Zaharchuk, S.K. Plevritis, Glioblastoma Multiforme: Exploratory Radiogenomic Analysis by Using Quantitative Image Features, Radiology 276 (1) (2015) 313.
[31] C.E. Ray Jr, R. Gupta, J. Blackwell, Changes in the American interventional radiology literature: comparison over a 10-year time period, Cardiovascular and interventional radiology 29 (4) (2006) 599–604.
[32] C.A. Karlo, P.L. Di Paolo, J. Chaim, A.A. Hakimi, I. Ostrovnaya, P. Russo, H. Hricak, R. Motzer, J.J. Hsieh, O. Akin, Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations, Radiology 270 (2) (2014) 464–471.
[33] S.L. Kerns, H. Ostrer, B.S. Rosenstein, Radiogenomics: using genetics to identify cancer patients at risk for development of adverse effects following radiotherapy, Cancer discovery 4 (2) (2014) 155–165.
[34] S. Bakr, O. Gevaert, S. Echegaray, K. Ayers, M. Zhou, M. Shafiq, H. Zheng, J.A. Benson, W. Zhang, A.N.C. Leung, M. Kadoch, C.D. Hoang, J. Shrager, A. Quon, D.L. Rubin, S.K. Plevritis, S. Napel, A radiogenomic dataset of non-small cell lung cancer, Scientific data 5 (2018) 180202.
[35] N.M. Braman, M. Etesami, P. Prasanna, C. Dubchuk, H. Gilmore, P. Tiwari, D. Plecha, A. Madabhushi, Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI, Breast cancer research : BCR 19 (1) (2017) 57.
[36] Z.A. Binder, A.H. Thorne, S. Bakas, E.P. Wileyto, M. Bilello, H. Akbari, S. Rathore, S.M. Ha, L. Zhang, C.J. Ferguson, S. Dahiya, W.L. Bi, D.A. Reardon, A. Idbaih, J. Felsberg, B. Hentschel, M. Weller, S.J. Bagley, J.J.D. Morrissette, M.P. Nasrallah, J. Ma, C. Zanca, A.M. Scott, L. Orellana, C. Davatzikos, F.B. Furnari, D.M. O'Rourke, Epidermal Growth Factor Receptor Extracellular Domain Mutations in Glioblastoma Present Opportunities for Clinical Imaging and Therapeutic Development, Cancer cell 34 (1) (2018) 163–177 e7.
[37] E. Rios Velazquez, C. Parmar, Y. Liu, T.P. Coroller, G. Cruz, O. Stringfield, Z. Ye, M. Makrigiorgos, F. Fennessy, R.H. Mak, R. Gillies, J. Quackenbush, H. Aerts, Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer, Cancer research 77 (14) (2017) 3922–3930.
[38] J. Yu, Z. Shi, Y. Lian, Z. Li, T. Liu, Y. Gao, Y. Wang, L. Chen, Y. Mao, Noninvasive IDH1 mutation estimation based on a quantitative radiomics approach for grade II glioma, European radiology 27 (8) (2017) 3509–3522.
[39] J.B. Permuth, D.T. Chen, S.J. Yoder, J. Li, A.T. Smith, J.W. Choi, J. Kim, Y. Balagurunathan, K. Jiang, D. Coppola, B.A. Centeno, J. Klapman, P. Hodul, F.A. Karreth, J.G. Trevino, N. Merchant, A. Magliocco, M.P. Malafa, R. Gillies, Linc-ing Circulating Long Non-coding RNAs to the Diagnosis and Malignant Prediction of Intraductal Papillary Mucinous Neoplasms of the Pancreas, Scientific reports 7 (1) (2017) 10484.
[40] F. Davnall, C.S. Yip, G. Ljungqvist, M. Selmi, F. Ng, B. Sanghera, B. Ganeshan, K.A. Miles, G.J. Cook, V. Goh, Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? Insights Imaging 3 (6) (2012) 573–589.
[41] V. Goh, B. Ganeshan, P. Nathan, J.K. Juttla, A. Vinayan, K.A. Miles, Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker, Radiology 261 (1) (2011) 165–171.
[42] B. Zhao, Y. Tan, W.Y. Tsai, J. Qi, C. Xie, L. Lu, L.H. Schwartz, Reproducibility of radiomics for deciphering tumor phenotype with imaging, Sci Rep 6 (2016) 23428.
[43] J.E. Park, S.Y. Park, H.J. Kim, H.S. Kim, Reproducibility and Generalizability in Radiomics Modeling: Possible Strategies in Radiologic and Statistical Perspectives, Korean J Radiol 20 (7) (2019) 1124–1137.
[44] P. Kalendralis, A. Traverso, Z. Shi, I. Zhovannik, R. Monshouwer, M.P.A. Starmans, S. Klein, E. Pfaehler, R. Boellaard, A. Dekker, L. Wee, Multicenter CT phantoms public dataset for radiomics reproducibility tests, Medical physics 46 (3) (2019) 1512–1518.
[45] M. Bogowicz, R.T.H. Leijenaar, S. Tanadini-Lang, O. Riesterer, M. Pruschy, G. Studer, J. Unkelbach, M. Guckenberger, E. Konukoglu, P. Lambin, Post-radiochemotherapy PET radiomics in head and neck cancer - The influence of radiomics implementation on the reproducibility of local control tumor models, Radiother Oncol 125 (3) (2017) 385–391.
[46] A. Chalkidou, M.J. O'Doherty, P.K. Marsden, False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review, PLoS One 10 (5) (2015) e0124165.
[47] R.B. Ger, C.E. Cardenas, B.M. Anderson, J. Yang, D.S. Mackin, L. Zhang, L.E. Court, Guidelines and Experience Using Imaging Biomarker Explorer (IBEX) for Radiomics, J Vis Exp (2018) 131.
[48] S.S. Hwang, H.H. Song, J.H. Baik, S.L. Jung, S.H. Park, K.H. Choi, Y.H. Park, Researcher contributions and fulfillment of ICMJE authorship criteria: analysis of author contribution lists in research articles with multiple authors published in radiology, International Committee of Medical Journal Editors, Radiology 226 (1) (2003) 16–23.