

REVIEW

Open Access



Imaging genomics of cancer: a bibliometric analysis and review

Xinyi Gou¹ , Aobo Feng², Caizhen Feng¹, Jin Cheng^{1*} and Nan Hong¹

Abstract

Background Imaging genomics is a burgeoning field that seeks to connections between medical imaging and genomic features. It has been widely applied to explore heterogeneity and predict responsiveness and disease progression in cancer. This review aims to assess current applications and advancements of imaging genomics in cancer.

Methods Literature on imaging genomics in cancer was retrieved and selected from PubMed, Web of Science, and Embase before July 2024. Detail information of articles, such as systems and imaging features, were extracted and analyzed. Citation information was extracted from Web of Science and Scopus. Additionally, a bibliometric analysis of the included studies was conducted using the Bibliometrix R package and VOSviewer.

Results A total of 370 articles were included in the study. The annual growth rate of articles on imaging genomics in cancer is 24.88%. China (133) and the USA (107) were the most productive countries. The top 2 keywords plus were “survival” and “classification”. The current research mainly focuses on the central nervous system (121) and the genitourinary system (110, including 44 breast cancer articles). Despite different systems utilizing different imaging modalities, more than half of the studies in each system employed radiomics features.

Conclusions Publication databases provide data support for imaging genomics research. The development of artificial intelligence algorithms, especially in feature extraction and model construction, has significantly advanced this field. It is conducive to enhancing the related-models’ interpretability. Nonetheless, challenges such as the sample size and the standardization of feature extraction and model construction must overcome. And the research trends revealed in this study will guide the development of imaging genomics in the future and contribute to more accurate cancer diagnosis and treatment in the clinic.

Advances in knowledge

- Imaging genomics is considered a crucial component in cancer precision medicine. A review is necessary to offer a holistic and systematic view of this field.
- Focused on survival and classification, imaging genomics shows potential in cancer heterogeneity, yet faces challenges on standardization, model complexity, and ethics.
- This study clarifies the objectives, current standing, and evolution of imaging genomics in cancer, offering a basic workflow and trends for cancer analysis in this domain.

*Correspondence:

Jin Cheng
chengjinpkuh@outlook.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Imaging genomics, Cancer, Bibliometric analysis, Review

Background

Cancer remains a devastating global health challenge, exerting a significant and escalating toll on societies and healthcare systems worldwide. The Global Cancer Statistics 2022 report reveals that in 2022, nearly 20 million individuals were diagnosed with cancer, resulting in 9.7 million cancer-related deaths [1].

Diagnostic imaging techniques such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and Ultrasound (US) are integral to the detection, staging, and monitoring of cancer treatment responses [2–4]. They can perform non-invasive multiple scans, providing information about the cancer and surrounding tissues, and can be compared for changes over time. With the advancement of artificial intelligence algorithms, the scope of imaging features has expanded beyond conventional imaging features, such as tumor size, shape, and margin, to encompass high-throughput, algorithm-based radiomics features [5, 6]. These mainly include handcrafted radiomic, and deep learning features, which may not be perceptible to the human eye [7].

In addition to imaging technology, advances in genomics have also provided new perspectives for cancer research, and the combination of the two has given birth to the interdisciplinary discipline of imaging genomics. Imaging genomics, also known as radiogenomics, is an emerging field that lies at the intersection of medical imaging and genomics [8, 9]. The study of genomics involves analyzing the complete genetic makeup, encompassing DNA examination, transcriptomics (which includes the investigation of mRNAs, miRNAs, and lncRNAs), and epigenomics research [10, 11]. The term radiogenomics, which first emerged in 2002, has expanded beyond its initial definition [12, 13]. Originally, it described research examining the relationship between patient genetics and variations in their sensitivity to radiation therapy. However, its content has now been expanded to include imaging genomics. The primary objective of imaging genomics is to identify and establish relationships between both image features (semantic and quantitative) with genomic information, thereby constructing association maps that can further be correlated with clinical outcomes or other relevant metrics [13, 14]. By creating a comprehensive map of associations, it can enhance the precision of prognostic and predictive models across a range of oncological conditions. This integration of imaging data with genomic insights holds the potential to improve patient stratification and treatment planning [15–17].

The underlying mechanism of imaging features can be discovered by imaging genomics, which may enhance the credibility of its clinical value and provide potential target of therapeutic intervention [18–20]. The formal classification of molecular subtypes through genomic analysis is both costly and time-consuming. In clinical practice, immunohistochemistry(IHC) methods are often employed as an alternative for molecular typing. Nevertheless, IHC specimens are typically extracted from a single localized region. Factors such as tissue preparation, tumor heterogeneity, or genetic variations affecting protein expression can result in discrepancies between IHC and genetic testing outcomes [21, 22]. Discordance rates between genetic testing and IHC classifications can indeed vary significantly, with reported rates ranging from about 40–100% across different studies [23–25]. Through imaging genomics, it becomes possible to predict molecular typing and therapeutic efficacy noninvasively.

Therefore, imaging genomics research has provided valuable insights into cancer [26]. This review synthesizes the current state of imaging genomics across various cancer types, focusing on their basic workflow and challenges. Additionally, a bibliometric analysis quantitatively assessed the trajectory and focal points of imaging genomics in cancer.

Materials and methods

Search strategy

Data collection was conducted utilizing the PubMed, Embase, and Web of Science databases, and was enhanced by examining the reference lists of pertinent articles included in our review up to July 2024. The search query of PubMed and Web of Science is (“imaging genomics” OR ‘radiogenomics’) AND “cancer”. For Embase, the search string is (‘imaging genomics’/exp OR ‘imaging genomics’ OR ‘radiogenomics’/exp OR ‘radiogenomics’) AND (‘cancer’/exp OR ‘cancer’). It is noteworthy that the initial objective of radiogenomics was to identify potential genomic predictors for radiation toxicity; however, this aspect was not included in the scope of the study.

Study selection

All articles were screened back-to-back by two independent investigators (XYG and ABF) to determine whether they met our inclusion criteria. Any disputes were solved by a discussion with a third reviewer (CZF). The inclusion criteria were as followed: (1) studies focusing on cancer patients; (2) Research utilizing imaging genomics methodologies; (3) Peer-reviewed and formally published

articles. And the exclusion criteria were: (1) non-human studies; (2) Articles not written in English; (3) Review.

Data extraction

Two investigators (XYG and ABF) extracted basic study information including author, year of publication, study design, sample size, and type of cancer. Subsequently, a more detailed extraction was performed to gather insights into the imaging genomics field, such as cancer types, imaging modalities, imaging features and genetic types. Imaging features included conventional imaging features and radiomics features.

Specifically, conventional imaging features refer to characteristics that can be directly observed, measured, and simply calculated from medical imaging, such as tumor size, shape, location, and signs (for example, mismatch on T2-FLAIR), as well as values obtained from special examinations like the apparent diffusion coefficient measurements from DWI sequences and the standardized uptake value max (SUVmax) values from PET scans. Radiomics features, on the other hand, are quantitative imaging characteristics derived from computer algorithms, such as handcrafted radiomics features based on the PyRadiomics platform and features extracted using a specific deep learning algorithm [27–29].

Bibliometric analysis

Citation data was extracted from the Web of Science and Scopus database to perform a bibliometric analysis. It involved the keywords, keywords plus, quantification of publication trends, citation patterns, and collaborative efforts within the imaging genomics of cancer research. Keywords plus is a set of terms extracted from the titles of cited documents in an article. It aims to complement the author-provided keywords, enhancing the thoroughness and precision of literature searches. Additionally, local citations refer to the frequency with which a study is referenced by other researchers within the included articles. Global Citations, on the other hand, indicate the total number of times an article has been cited across all documents indexed in extensive bibliographic databases, such as Web of Science or Scopus. The data will be used to identify key publications and keywords in the research that require further investigation. Bibliometrix R package(<https://github.com/massimoaria/bibliometrix>) and VOSviewer were used to perform bibliometric analyses and visualization.

Results

Published literature was retrieved from PubMed ($n = 572$), Embase ($n = 1123$) and Web of Science ($n = 964$). Based on the inclusion and exclusion criteria, a total of 370 articles were selected for this review (Fig. 1).

Annual publication output

From 2008 to June 2024, a total of 370 articles were published from 151 journals on the topic of cancer and imaging genomics. The data reveals a significant compound annual growth rate of 24.88%, indicating a rapidly increasing interest and research activity in this field. The internal co-authorship rate was 26.49%, with an average of 10.2 co-authors per document. A total of 10,528 references were cited across these publications.

Affiliation and country

As shown in Fig. 2A, the most prolific source of published articles is the University of California, with a total of 74 publications. In second place, the University of Texas boasts 60 articles. Sharing the third position are the Memorial Sloan Kettering Cancer Center and Stanford University, both contributing 50 articles each. Fudan University secures the fifth spot with 45 articles, closely followed by the Helmholtz Association with 44 articles. From the data presented, it is clear that the University of California leads in research output. This likely reflects the institution's strong research capabilities and extensive academic resources in the relevant fields.

After an analysis of corresponding author's countries, it was exhibited that a total of 27 countries/regions included articles. As shown in Fig. 2B, the country that contributed the largest volume of publications (133) was the China, which accounted for 35.9% of the total. USA was ranked the second (107, 28.9%), Korea third (22, 5.9%), and followed by Italy(19, 5.1%) and Germany (17, 4.6%). Multiple country publications (MCP) revealed the publication number of co-authors from different countries/regions. Although the USA had the highest MCP (35), its MCP ratio ($= \text{MCP/articles}$) was 32.7%. The top 3 cited countries were USA(5194), Netherlands(3336) and China(2339). Countries with close cooperation include USA and CHINA (22), USA and Germany (10), USA and United Kingdom (9). This indicates that the United States plays a significant role in international scientific research cooperation and maintains particularly frequent collaborative relationships with China. The collaboration between the two countries likely facilitates the exchange of research resources and technology, thereby enhancing their respective research capabilities.

Source journals

In the bibliometric analysis of primary source journals, "Cancers" has emerged as the most prolific journal, with a total of 28 articles, underscoring its pivotal role in oncology research publication. It is closely followed by "Scientific Reports", which has contributed a notable 20 articles, reflecting its multidisciplinary research focus. Ranking third is "European Radiology", which has contributed 19 articles, highlighting its substantial impact in

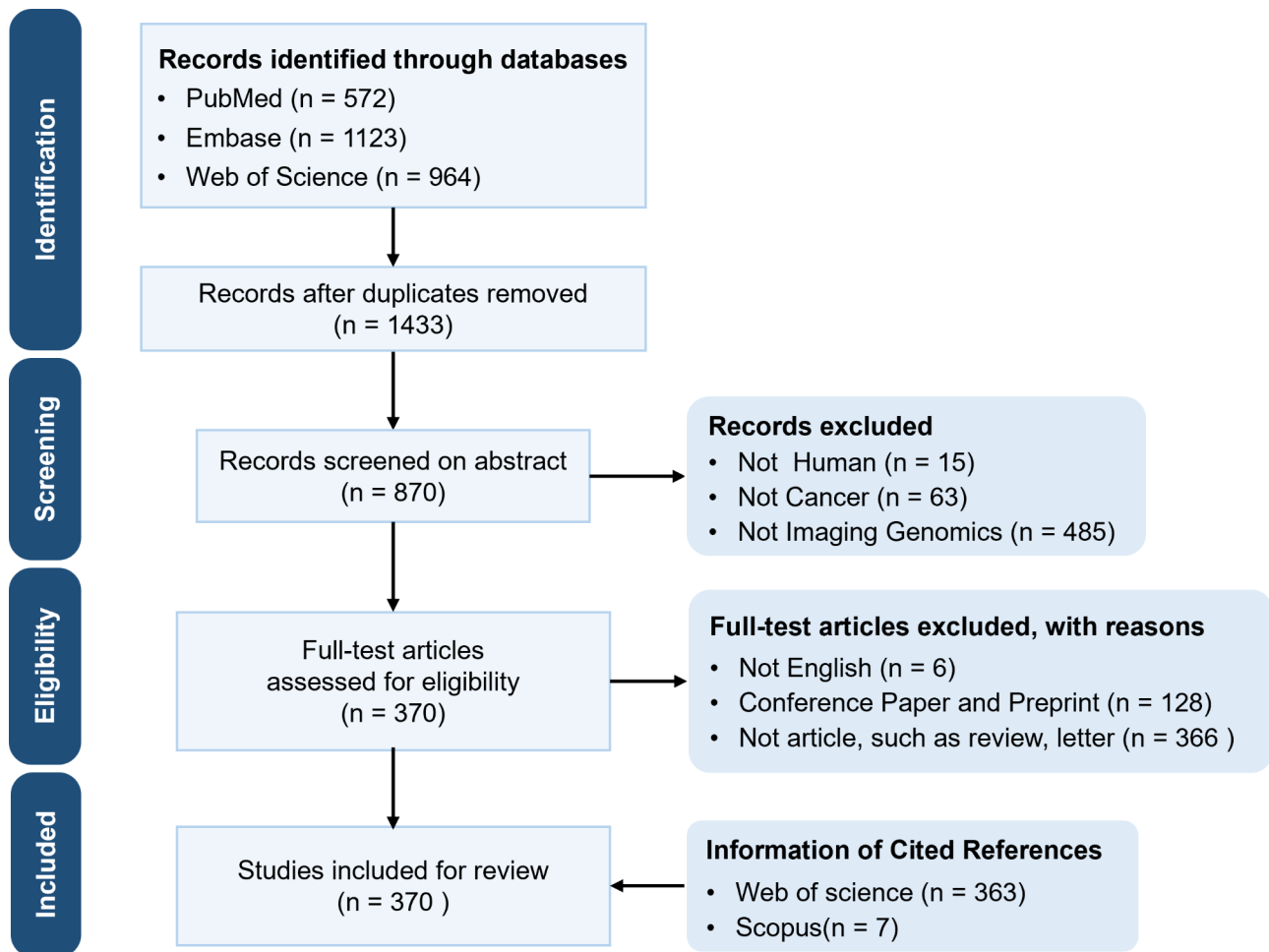


Fig. 1 Schematic workflow of this review

the field of medical imaging. “Frontiers in Oncology” and “Radiology” are tied for the fourth position, each journal contributing 16 articles, respectively.

In terms of the most locally cited sources, “Radiology” leads with an impressive 885 citations. “European Radiology” and “Neuro-Oncology” are the second most cited with 330 citations, followed by “Clinical Cancer Research” with 329 citations, and “Scientific Reports” with 314 citations. The most globally cited document is a paper authored by Hugo J. W. L. Aerts and colleagues, which was published in the journal *Nature Communications* in June 2014 and has garnered a significant global citation count of 3,281 (Table 1) [30]. This research represents an early application of imaging genomics, examining CT scans from 1019 patients with lung and head and neck cancers across seven distinct cohorts [31]. The study aims to identify radiological characteristics that reflect intra-tumor heterogeneity and investigate the potential gene expression patterns associated with these features. As for the most locally cited document, it is a publication

by GEVAERT O in the journal *Radiology* from 2012, with 33 local citations and a total of 310 global citations [32].

Keyword occurrence analysis

A comprehensive collection of 919 unique keywords plus was assembled. Figure 3A illustrates the distribution of these keywords plus using a treemap visualization. The keyword plus “Survival” predominated with the highest frequency, appearing 79 times, indicating its significant role in the research field. It was closely trailed by “Classification” with 63 occurrences. These two high-frequency terms are interconnected and reflect the current research priorities.

Among the top 20 most frequently used keywords, ‘Glioblastoma’ was noted with 20 occurrences, standing out as the sole cancer type to make it onto the list, highlighting its particular relevance to the current body of research.

Moreover, Fig. 3B showed the citation network map of the top 20 keywords plus. The keywords plus of high-citation articles were as followed: “heterogeneity”,

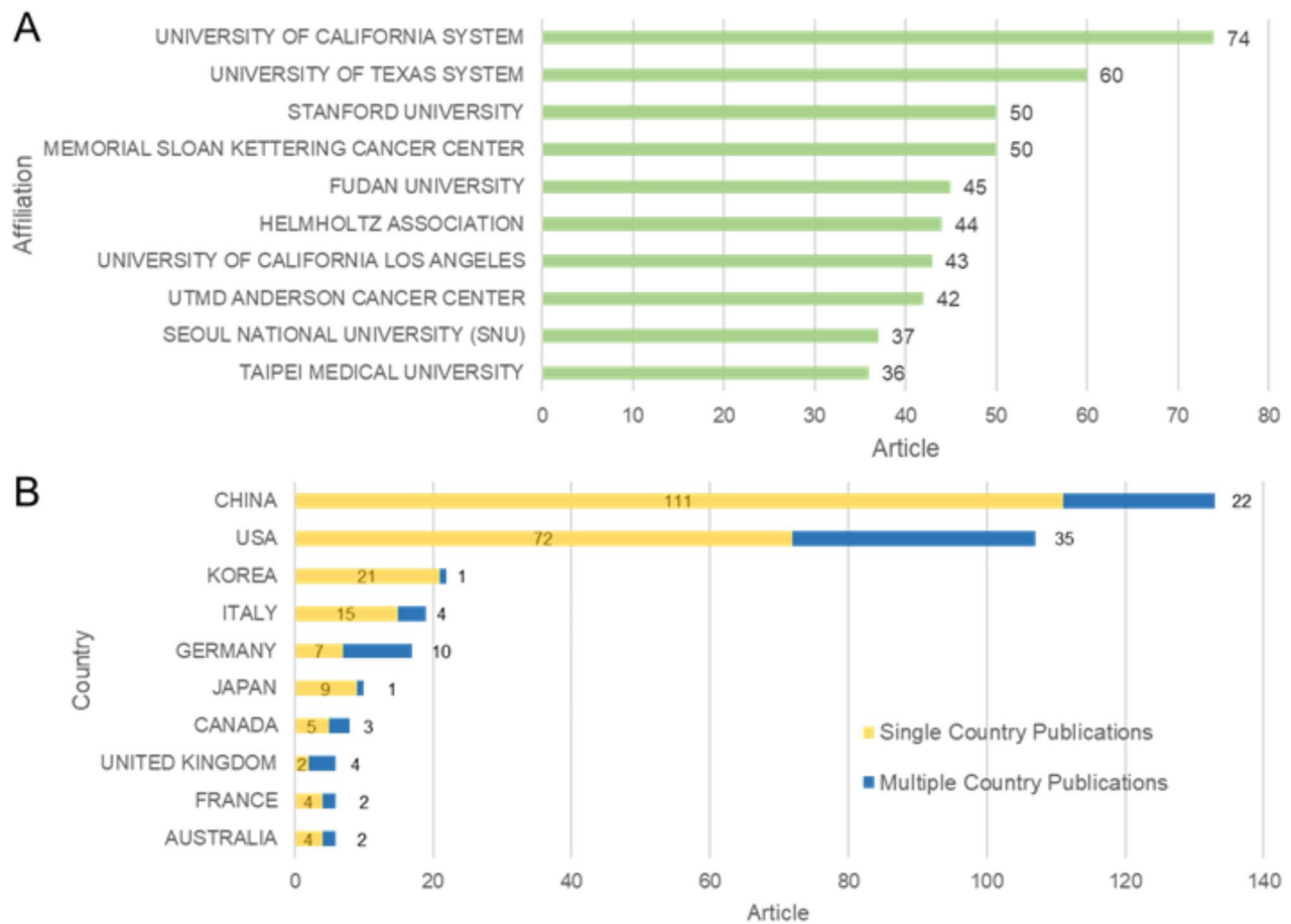


Fig. 2 **A**, Bar chart showing the number of articles published by the top 10 Affiliations; **B**, Bar chart displaying the number of articles published by the top 10 Countries, with SCP representing Single Country Publications and MCP representing Multiple Country Publications

“gene-expression”, “temozolomide”, “prediction”, “mutations” and “chemotherapy”. These words indicate that current research primarily focuses on improving disease prediction accuracy, understanding tumor heterogeneity, exploring related gene expression patterns, and evaluating drug efficacy to optimize chemotherapy strategies. Among these, ‘temozolomide’ is a chemotherapy drug primarily used in the treatment of certain types of brain tumors. It achieves its therapeutic effect by methylating DNA, which inhibits the cells’ ability to replicate and grow.

We conducted keywords and publication year analyses related to the central nervous system, genitourinary system, respiratory system, and digestive system (Fig. 4). In recent years, the main keyword was “deep learning” in the central nervous system, genitourinary system was “prognosis”, “diagnosis”, “machine learning”, “computed tomography”, respiratory system was “non-small cell lung cancer”, “cell lung cancer”, “features”, “egfr mutation”, “machine learning”, “heterogeneity”, and digestive system was “radiomics”, “hepatocellular carcinoma”, “MRI” and “signatures”. These keywords highlight the primary

focuses within each system. “Deep learning” has become central in nervous system research, reflecting artificial intelligence’s rise in neuroimaging. Researchers use deep learning algorithms to enhance disease diagnosis, prognosis prediction, and treatment outcome evaluation. This technology improves accuracy and speeds up data extraction, offering tools for personalized medicine. In the genitourinary system, research aims to improve prognostic assessment and diagnostic precision. Combining “machine learning” and “computed tomography” enables more accurate lesion identification and personalized treatment plans. Predicting disease progression and treatment response remains a key focus. Respiratory system research emphasizes non-small cell lung cancer, particularly EGFR gene mutations. “Machine learning” technology helps researchers understand tumor heterogeneity, crucial for precise treatment strategies. In digestive system research, radiomics and hepatocellular carcinoma are prominent topics. Radiomics involves quantitative analysis of imaging data, revealing tumor structure and behavior. MRI supports precise diagnostics in the digestive system.

Table 1 The top 10 most globally cited documents in the imaging genomics

Paper	DOI	Total Citations	Total Citations per Year
AERTS H, 2014, NAT COMMUN	https://doi.org/10.1038/ncomms5006	3281	273.42
WANG S, 2019, EUR RESPIR J	https://doi.org/10.1183/13993003.00986-2018	274	39.14
GEVAERT O, 2012, RADIOLOGY	https://doi.org/10.1148/radiol.12111607	310	22.14
KICKINGEREDER P, 2015, SCI REP-UK	https://doi.org/10.1038/srep16238	231	21.00
KICKINGEREDER P, 2016, RADIOLOGY	https://doi.org/10.1148/radiol.2016161382	208	20.80
DIEHN M, 2008, P NATL ACAD SCI USA	https://doi.org/10.1073/pnas.0801279105	352	19.56
BUDA M, 2019, COM-PUT BIOL MED	https://doi.org/10.1016/j.compbiomed.2019.05.002	136	19.43
GEVAERT O, 2014, RADIOLOGY	https://doi.org/10.1148/radiol.14131731	230	19.17
KARLO C, 2014, RADIOLOGY	https://doi.org/10.1148/radiol.13130663	213	17.75
LIU X, 2018, NEUROIMAGE-CLIN	https://doi.org/10.1016/j.nicl.2018.10.014	142	17.75

The central Nervous System category included 121 articles, the genitourinary system had 110 (with breast cancer comprising 44 of these), the respiratory system accounted for 68 articles, the digestive system contributed 35, and the other systems category encompassed 36 articles (Fig. 5). The imaging modalities involved were as follows: 134 articles utilized CT, 193 employed MR, 28 used PET, 6 incorporated US, 3 applied mammography, and the remaining 18 articles featured multiple modalities. As depicted in Fig. 5, there is considerable variation in the imaging modalities used across different systems within imaging genomics. Notably, research pertaining to cancers of the central nervous system predominantly utilizes MRI (98.4%), whereas studies on cancers of the respiratory and digestive systems primarily employ CT scans. This disparity may be attributed to the development and application of imaging scanning technologies in different system.

Moreover, radiomics features have been the focus of 237 articles, taking the lead. Conventional imaging features are represented in 110 articles, and their combined application is discussed in 35 articles. Across various systems, articles utilizing radiomics features outnumber those employing conventional imaging features. In addition, studies examining gene expression and function were comparable in number to those investigating gene mutation and modification in both the respiratory and digestive systems.

Discussion

Current status of the research

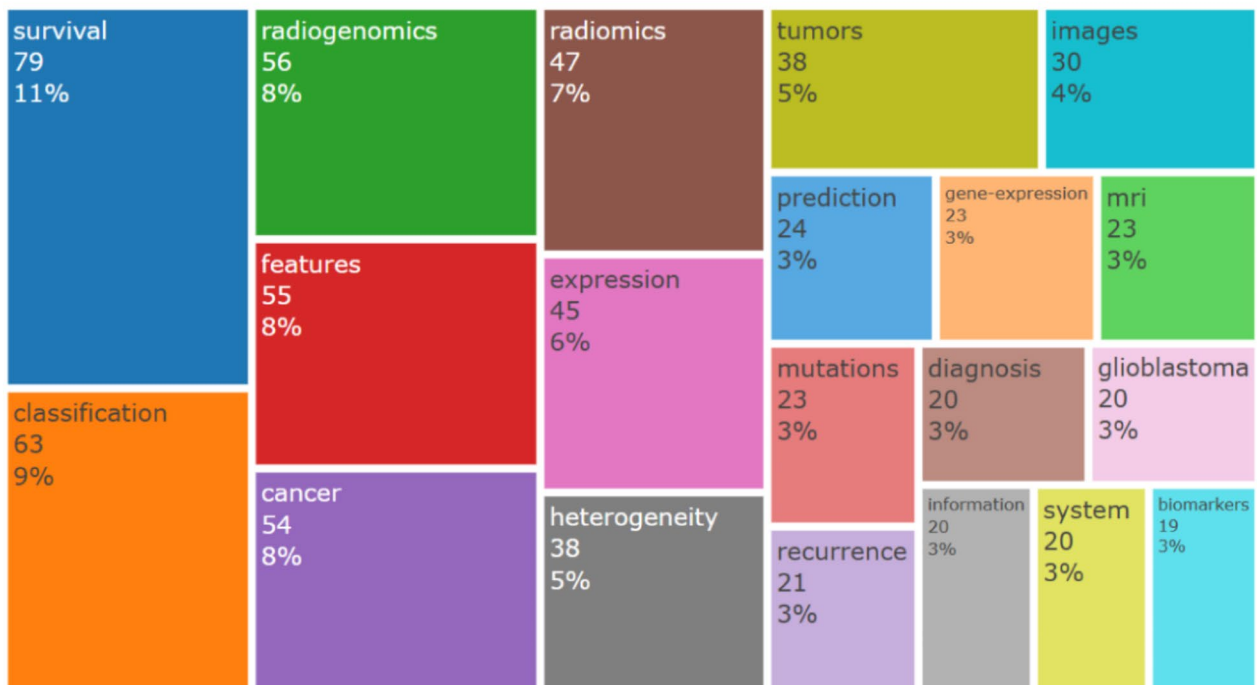
The study analysis 370 articles about imaging genomics in cancer, which were from 2008 to 2024.

Imaging genomics, referred to radiogenomics, has become increasingly prevalent in cancer-related research, with a significant upward trend in publication volume over the year. On the one hand, public databases such as The Cancer Genome Atlas (TCGA) and The Cancer Imaging Archive (TCIA) provide crucial data support for the advancement of imaging genomics [33–35]. On the other hand, the advancement of artificial intelligence algorithms has enabled the extraction of high-throughput radiomics features and gene-related features [36–39].

This study meticulously analyzed and synthesized data from 370 included articles, which were from 2008 to 2024. The data revealed that the research has predominantly been concentrated in the United States and China, with close collaboration between the two nations, which is attributed to their innovative analytical techniques and extensive patient cohorts. Additionally, the research mainly focus on gliomas and breast cancer in this field [26, 33]. It may be because these cancer types have been extensively studied at the genomic level, and there is a growing body of evidence linking specific genomic alterations to imaging features. For gliomas, the emergence of novel advanced imaging techniques and the extensive work done by TCGA and the Ivy Glioblastoma Atlas Project on mapping genomic changes have led to the discovery of new correlations between genomic alterations and imaging features [40, 41]. Breast cancer, as the second common cancer type, has been found to be clinically and genomically heterogeneous, and incorporating genomic information into treatment decisions is a promising area of focus [1, 42, 43].

Bibliometric analysis reveals that “survival” and “classification” are the predominant keywords plus, underscoring the central themes in cancer imaging genomics research. “Survival” typically focuses on identifying features that may influence the prognosis of cancer patients, thereby further exploring related molecular mechanisms and predicting patient survival as well as recurrence and metastasis. A recent study analyzed multi-omics data from 1,754 glioma patients, identifying 35 genes and 82 radiomic features highly correlated [44]. This led to an effective predictive model for mortality risk and revealed associations with the tumor immune microenvironment. It indicates that genomics-imaging interaction can aid in accurately predicting glioma prognosis and immune responses. “Classification” centers on patient status, such as molecular subtyping and response to therapy. It involves investigating differences in imaging and/or gene features across different patient groups to explore the heterogeneity of cancers. It aids in establishing predictive

A



B

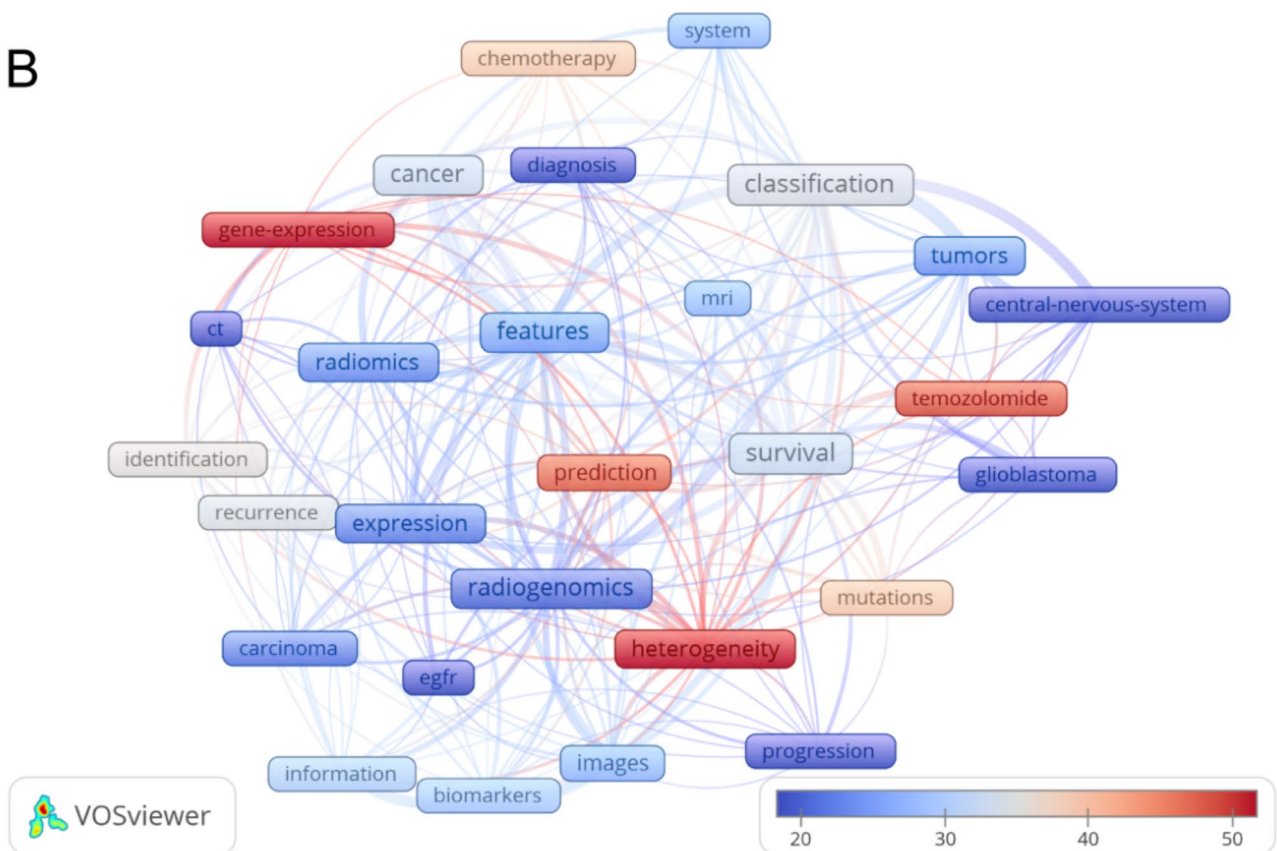


Fig. 3 **A**, Treemap of top 20 Keywords plus; **B**, the average citation network map of the 28 keywords plus (node size represents the frequency of keyword occurrence in the included literature, and color indicates the average number of citations of documents associated with keywords)

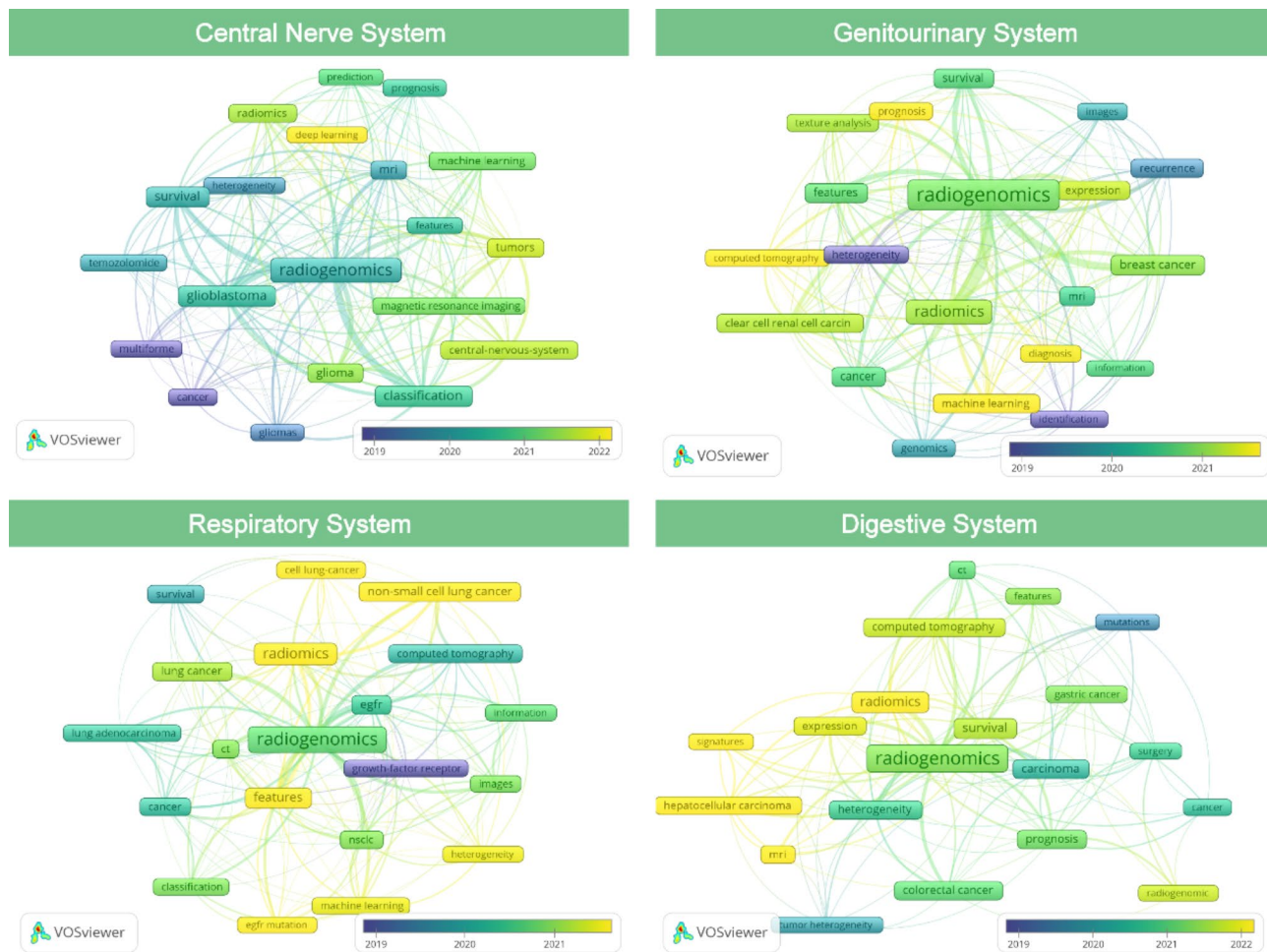


Fig. 4 The network map of keywords in four main systems (node size represents the frequency of keyword occurrence in the included literature, and color indicates the publication years)

and evaluative models, which can facilitate personalized medicine by tailoring treatments to individual patient characteristics. Furthermore, the primary objectives of current research were as follows: 1) To investigate the relationship between imaging features and gene features; 2) To develop molecular typing and associated predictive models; 3) To establish models for predicting and assessing treatment outcomes; 4) To create risk stratification and prognostic prediction models.

Basic workflow

In cancer research, the fundamental workflow of imaging genomics can be segmented into four key components: data acquisition, imaging and gene features extraction, correlation analysis, and model construction (Fig. 6).

Imaging features

Imaging data can be provided by hospitals, institutions or public databases. The preeminent database for imaging genomics research is the Cancer Imaging Archive (TCIA), primarily because it merges TCGA to provide a

comprehensive collection of clinical, genetic, and pathological information. Genomic data from TCGA is typically provided in several formats, including raw counts, TPMs (Transcripts Per Million), and FPKM (Fragments Per Kilobase of transcript per Million mapped reads). Moreover, the TCGA platform offers detailed information on mutation statuses and conducts foundational differential gene expression analysis. Common cancer imaging databases have been summarized by Piotr Woznicki et al. and Yusheng Guo et al. [35, 45].

Imaging features can be extracted after data preprocessing and segmentation. It can be categorized into two main types: conventional features (encompassing gross imaging features and quantitative imaging parameters) and radiomics features (which include both hand-craft radiomics features and deep learning features). This study found that more than half of the included articles used radiomics features, which may be because radiomics features are large and relatively easy to obtain. Junior physicians can independently and relatively accurately outline lesions and tissues after a short training

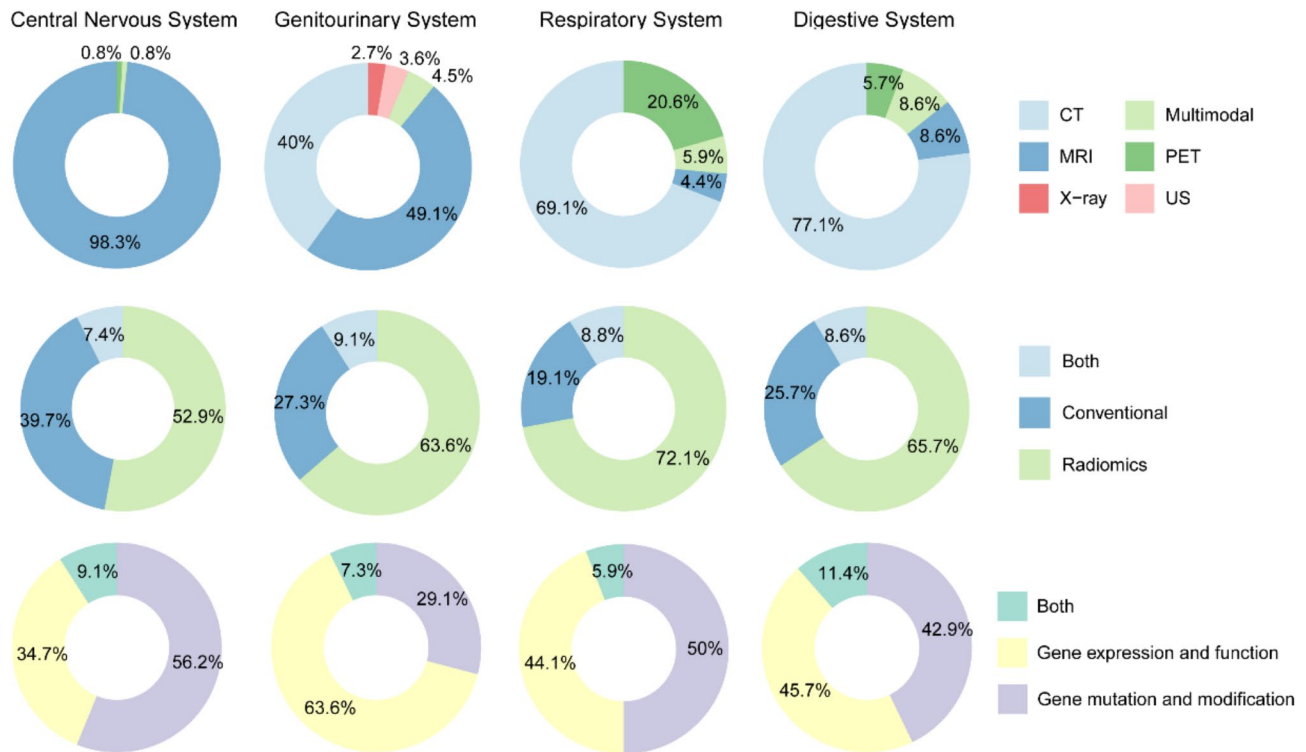


Fig. 5 Pie charts illustrating the distribution of proportions for various imaging modalities, image features, and gene features across four major systems as reported in the included literature

period. A large number of handcraft radiomics are available through Pyradiomics, HeterogeneityCAD or QIFE (Quantitative Imaging Feature Extraction) [46, 47]. They include first-order features(basic statistics of pixel intensity, second-order features(texture measures) and higher-order features. Moreover, select an appropriate deep learning framework to construct the deep learning model. Upon successful training of the model, one can proceed to extract deep learning features from different feature maps. These features are typically derived from intermediate layers within the model, especially from bottleneck layers, which are critical points where the network compresses and encodes essential information about the input data into a more abstract and condensed representation.

While some conventional features are obtained based on the rich clinical experience and summary validation of radiologists, the accurate assessment of some conventional features relies on the experience of radiologists, such as extramural vascular invasion score [48]. Because of the heterogeneity of cancer, prediction and assessment of molecular typing of cancer and its prognosis often requires more features. However, extracting a number of conventional features by senior radiologists is relatively time-consuming.

Gene features

Gene data can be obtained by whole genome sequencing, gene expression microarrays or quantitative PCR [49]. Currently, some gene features based on single-cell sequencing have also been applied to imaging genomics [45]. Common cancer-related gene databases include TCGA, Gene Expression Omnibus (GEO), and cBioPortal, which have DNA, transcriptomics, epigenomics or proteomics.

Imaging genomics research typically concentrates on analyzing gene expression, functionality, mutation, and modifications. The biological behavior of cancer is often too complex to be fully captured by the expression or mutation of a single gene [50]. Consequently, some imaging genomics research in cancer typically investigates alterations in multiple genes. In breast cancer, researchers frequently emphasize the OncotypeDX:21-gene Recurrence Score Assay, which can provide crucial insights into both the risk of recurrence and the potential efficacy of chemotherapy [51, 52]. Various analytical methods, including Gene Ontology, Kyoto Encyclopedia of Genes and Genomes, and Gene Set Enrichment Analysis, can be utilized to examine the functionality of genes or gene clusters. CIBERSORT is a computational tool designed for the analysis of immune cell composition in complex tissues using high-dimensional genomic data, particularly gene expression profiles from bulk tissue

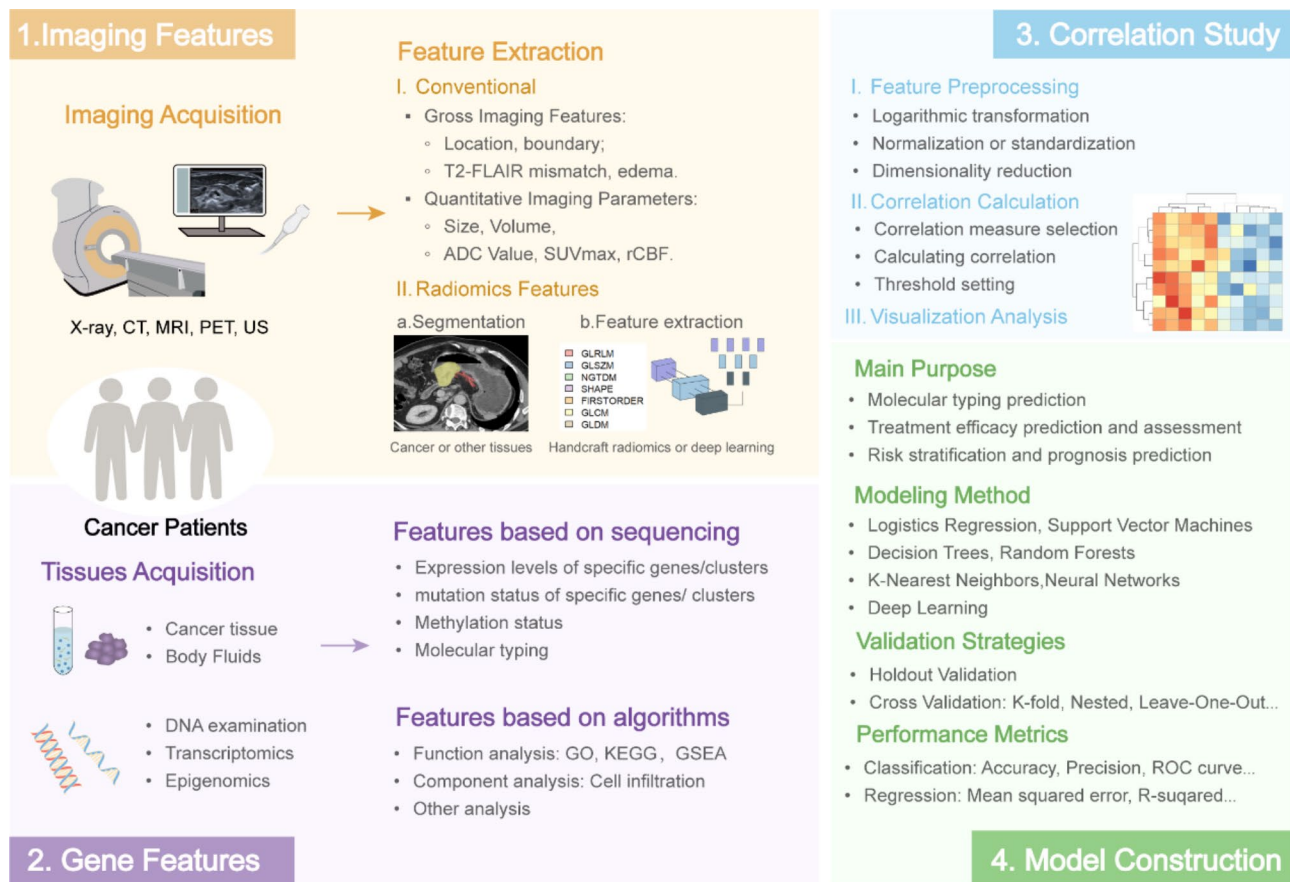


Fig. 6 The basic workflow of Imaging genomics in cancer research

samples [53]. By estimating the abundance of immune and stromal cell populations in heterogeneous tissues, it can be used to estimate the relationship between cell subpopulations and treatment response and survival outcomes in cancer [54–56]. Furthermore, due to the transformative phase in genomics research, characterized by advanced bioinformatics technologies and AI integration, an increased amount of genetic and other biological information can now be extracted and analyzed with improved accuracy [45, 57].

Correlation study

In the field of imaging genomics research, the most prevalent correlation study involves identifying imaging features associated with genetic mutations, gene expression, or molecular tumor subtypes. Concurrently, certain studies investigate the molecular mechanisms underlying the appearance of imaging features. The research conducted by Yang Hao et al. examines the molecular mechanisms of CT-detected extramural venous invasion in gastric cancer, as well as its relationship with macrophages and cancer associated fibroblasts in the tumor microenvironment [58–60]. While the present study focuses on the utilization of radiomics features, it is conceivable that

incorporating genomic analysis could enhance the understanding of the complex, high-dimensional radiomics features currently. Furthermore, some studies aim to identify image and genomic features associated with cancer patient prognosis or treatment efficacy [61–63]. These studies attempt to develop models using both image and genomic features. However, datasets containing paired image and gene information are typically limited in size.

Model construction

When assessing correlations and constructing models, it is paramount to tailor the analytical approach to the specific goals of the research and the distinctive attributes of the involved features. Deep learning models, such as Convolutional Neural Networks, demonstrate superior accuracy and performance in handling complex imaging tasks [64]. However, these models require large datasets to prevent overfitting and have high demands on computational resources, while their decision-making process is often opaque, which can be a drawback in clinical settings where transparency is crucial [43–45]. In contrast, simple models like logistic regression or decision trees perform well when data is scarce or computational resources are limited, and they offer greater interpretability.

However, they are less effective at capturing complex patterns in medical images. Therefore, model selection should account for task complexity, data volume, computational power, and interpretability.

Challenges and future directions

Numerous studies have demonstrated the exceptional performances of these models. However, several challenges remain when considering the transition from theoretical models to practical clinical applications.

Ethical considerations of data

In the field of imaging genomics research, concerns about privacy and data protection are fundamental and highly significant [65, 66]. Organizations must navigate a complex landscape of regulations including General Data Protection Regulation and Health Insurance Portability and Accountability Act, which govern the use and sharing of genomic and imaging data. Adrien Oliva and colleagues have also outlined approaches and recommendations regarding the storage of genomic and health care data, privacy and security measures, and the management of informed consent [67].

Emerging technologies such as artificial intelligence (AI) raise additional ethical questions regarding bias in algorithms used for analysis or decision-making. This bias can stem from various sources, including data inequities and modeling choices. The development and validation processes of AI algorithms require transparency. Establishing clear definitions for fairness and accountability in AI systems is equally vital. As a result, relevant regulatory frameworks should adapt to the changing landscape and offer more comprehensive governance solutions.

Standardization in the research process

The challenge of feature standardization is well-documented, particularly within the realm of radiomics. Recently, guidelines such as the CheckList for Evaluation of Radiomics research (CLEAR) and the Radiomics Quality Score (RQS) approach have been published with the aim of improving the quality and rigor of radiomic research [8, 66, 68]. They covered study design, ethical considerations, data management, segmentation, pre-processing, feature extraction, and result dissemination. The variability in imaging equipment, including differences across manufacturers and scanning protocols, can influence imaging features [69]. As a result, data pre-processing is essential, and the selection of methods and parameter settings should be determined based on the specific condition. Certain researchers have compiled relevant guidelines [70, 71]. And the CLEAR checklist also recommends that studies related to radiomics should provide details on image preprocessing techniques [66].

In additions, to improve the consistency of conventional imaging feature extraction, the open-source ePAD platform (<https://epad.stanford.edu>) has been developed to facilitate quantitative imaging feature annotation and sharing [72]. For the consistency of radiomics feature extraction, medical imaging auto-segment is useful [73].

Similarly, gene features are susceptible to the impact of detection methodologies and bioinformatic analysis. The Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists have jointly published the “Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer,” which provides a consensus recommendation for the analysis and communication of genetic variations found in cancer contexts [74]. Moreover, Christina A. Austin-Tse and her colleagues contribute to the development of recommended protocols for clinical whole genome sequencing [75]. Additionally, certain analysis platforms offer standardized workflows. One such example is RaNA-Seq (<https://ranaseq.eu/home>), an accessible bioinformatics tool designed for swift RNA-Seq data analysis. This platform processes FASTQ files, generates quality control metrics, performs differential expression analyses, and aids in functional interpretation.

Model performance improvement

The model bias is mainly related to the characteristics and number of datasets used to train the model. Therefore, for Imaging genomics research, it is essential to have detailed patient inclusion and exclusion criteria and patient clinical characteristics. A large sample size is key to improving the model effectiveness and generalization ability. A large sample size required a large amount of manual segmentation work at the beginning. Existing automatic segmentation tools, such as MedSAM, MedSeg or MONAI, can be considered [76, 77]. However, these tools primarily focus on tissue or organ segmentation. Segmenting a part of the sample manually and training the automatic segmentation model to complete the subsequent segmentation can also be considered. To ensure accuracy, Dice scores need to be used to compare with the function obtained after manual segmentation, and radiologists also still need to check segmentation on a case-by-case basis [70]. And the establishment of databases containing medical segmentation, predictive modeling, etc. is important for the clinical translation of research results.

In addition, for most cases where sufficient data is not available, it is crucial to explore several strategies to enhance data availability and improve model performance. This includes fostering multi-center cooperation, which allows different institutions to collaborate and share their datasets. Furthermore, making full use of public databases can provide valuable supplementary

datasets that may help fill gaps in local data availability. Employing sensible data augmentation techniques can also significantly enhance the diversity of the training dataset by artificially expanding it with variations of existing data. This approach helps models generalize better by exposing them to a wider range of scenarios. Researchers can also enhance deep learning model performance on limited datasets by using transfer learning and regularization methods [78–80]. Additionally, leveraging multimodal data—such as combining different imaging modality, gene-relevant information, proteomics and metabolomics—can provide a more comprehensive understanding. Similarly, incorporating time-series data can also capture dynamic changes in patient conditions over time, further enriching the dataset. Memorial Sloan Kettering Cancer Center has developed MSK-CHORD, a comprehensive dataset integrating natural language processing annotations with structured data on medications, demographics, and genomics from over 24,950 patients across various cancer types [81]. Machine learning models trained on this dataset have shown superior performance in predicting survival, and analysis of radiology reports within MSK-CHORD has revealed predictors of metastasis to specific organ. It demonstrated the significance of automated annotation of clinical notes and data integration to enhance predictions of patient outcomes.

AI integration for clinical translation

AI drives the advancement of imaging genomics in four key areas: automating the extraction of high-dimensional imaging and genomic features, facilitating the integration of multi-omics data, offering advanced model architectures, and enhancing the visualization of high-dimensional data to improve model interpretability.

As the results of this study show, the imaging features of existing studies are mainly radiomics features. Reducing time-consuming repetitive work through automated image preprocessing and automatic segmentation based on AI, completing radiomics feature extraction, and evaluating and predicting based on the constructed model will help accelerate its clinical application. Meanwhile, automated feature extraction also helps to reduce the subjective bias of manual outlining. And the enhanced interpretability of the model through gene-based validation and AI visualization algorithms (such as SHapley Additive exPlanations and Gradient-weighted Class Activation Mapping) will also help physicians understand the decision path of the model and enhance the acceptance of its application [82–84].

Conclusions

Imaging genomics has made significant strides in cancer research, with a notable increase in publications and a focus on diseases like gliomas and breast cancer. Despite

advancements, challenges in feature standardization and data variability persist, necessitating solutions like the CLEAR checklist and RQS for quality assurance. Moving forward, the field is set to benefit from enhanced collaboration, larger datasets, and the continued evolution of artificial intelligence in feature extraction.

Abbreviations

AI	Artificial intelligence
CT	Computed Tomography
DWI	Diffusion Weighted Imaging
GEO	Gene Expression Omnibus
IHC	Immunohistochemistry
MCP	Multiple country publications
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
RQS	Radiomics Quality Score
SUVmax	Standardized Uptake Value max
TCGA	The Cancer Genome Atlas
TCIA	The Cancer Imaging Archive
TPM	Transcripts Per Million
US	Ultrasound

Acknowledgements

We thank all authors for their contribution in imaging genomics.

Author contributions

All authors read and approved the final manuscript. Xinyi Gou: Conceptualization, Methodology, Data curation, Writing-Original draft preparation, Writing-Reviewing and Editing. Aobo Feng: Methodology, Investigation, Data curation, Writing-Reviewing and Editing. Caizhen Feng: Methodology, Investigation, Data curation, Writing-Reviewing and Editing. Jin Cheng: Conceptualization, Administrative support, Investigation, Supervision, Writing-Reviewing and Editing. Nan Hong: Administrative support, Supervision, Writing-Reviewing.

Funding

This work was supported by Peking University People's Hospital Scientific Research Development Funds under Grant Number RDX2024-08.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Radiology, Peking University People's Hospital, Beijing, China

²College of Computer and Information, Inner Mongolia Medical University, Inner Mongolia, China

Received: 9 December 2024 / Accepted: 13 February 2025

Published online: 04 March 2025

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and

- mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229–63.
2. Panayides AS, Pattichis MS, Leandrou S, Pitris C, Constantinidou A, Pattichis CS. Radiogenomics for Precision Medicine with a Big Data Analytics Perspective. *IEEE J Biomedical Health Informat*. 2019;23(5):2063–79.
 3. Korn R. Novel multi-modal endpoints uncovered by radiogenomics integrating CT or MRI data. *Ann Oncol*. 2015;26(Korn R Imaging Endpoints Core Lab, Scottsdale, AZ, United States):ii13.
 4. Garcia-Figueiras R, Baleato-Gonzalez S, Luna A, Munoz-Iglesias J, Oleaga L, Vallejo Casas JA, et al. Assessing immunotherapy with functional and molecular imaging and Radiomics. *Radiographics*. 2020;40(7):1987–2010.
 5. Ferro M, de Cobelli O, Musi G, del Giudice F, Carrieri G, Busetto GM, et al. Radiomics in prostate cancer: an up-to-date review. *Ther Adv Urol*. 2022;14:17562872221109020.
 6. Mayerhoefer ME, Materka A, Langa G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. *J Nucl Med*. 2020;61(4):488–95.
 7. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, et al. Deep learning: a primer for radiologists. *Radiographics*. 2017;37(7):2113–31.
 8. Singh S, Mohajer B, Wells SA, Garg T, Hanneman K, Takahashi T, et al. Imaging Genomics and Multiomics: a guide for beginners starting Radiomics-Based Research. *Acad Radiol*. 2024;31(6):2281–91.
 9. Hartmann K, Sadée CY, Satwah I, Carrillo-Perez F, Gevaert O. Imaging genomics: data fusion in uncovering disease heritability. *Trends Mol Med*. 2023;29(2):141–51.
 10. Menghi F, Liu ET. Functional genomics of complex cancer genomes. *Nat Commun*. 2022;13(1):5908.
 11. Onoyama T, Ishikawa S, Isomoto H. Gastric cancer and genomics: review of literature. *J Gastroenterol*. 2022;57(8):505–16.
 12. Yard BD, Adams DJ, Chie EK, Tamayo P, Battaglia JS, Gopal P, et al. A genetic basis for the variation in the vulnerability of cancer to DNA damage. *Nat Commun*. 2016;7:11428.
 13. Pinker K, Shitano F, Sala E, Do RK, Young RJ, Wibmer AG, et al. Background, current role and potential applications of radiogenomics. *J Magn Reson Imaging*. 2018;47(3):604–20.
 14. Liu Z, Wu K, Wu B, Tang X, Yuan H, Pang H, Imaging genomics for accurate diagnosis and treatment of tumors: A cutting edge overview. *Biomed Pharmacother*. Wu Z, Wu K, Tang B, Yuan X, Pang H, Huang H, Zhu Y et al. X, xzhu@gdmu.edu.cn; Luo H, luohui@gdmu.edu.cn; Qi Y, qiyi7272@gdmu.edu.cn Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang), Zhanjiang, China:111173.
 15. Bodalal Z, Trebeschi S, Nguyen-Kim TDL, Schats W, Beets-Tan R. Radiogenomics: bridging imaging and genomics. *Abdom Radiol (NY)*. 2019;44(6):1960–84.
 16. Wang Y, Wang Y, Guo C, Xie X, Liang S, Zhang R, et al. Cancer genotypes prediction and associations analysis from imaging phenotypes: a survey on radiogenomics. *Biomark Med*. 2020;14(12):1151–64.
 17. Mendes Serrão E, Klug M, Moloney BM, Jhaveri A, Lo Gullo R, Pinker K, et al. Current status of Cancer Genomics and Imaging phenotypes: what radiologists need to know. *Radiol Imaging Cancer*. 2023;5(6):e220153.
 18. Gao B, Feng C, Chai F, Wei S, Hong N, Ye Y, et al. CT-detected extramural venous invasion-related gene signature for the overall survival prediction in patients with gastric cancer. *Cancer Med*. 2021;10(21):7816–30.
 19. Ogbonnaya CN, Alsaedi BSO, Alhussaini AJ, Hislop R, Pratt N, Nabi G. Radiogenomics Reveals Correlation between Quantitative Texture Radiomic Features of Biparametric MRI and Hypoxia-Related Gene Expression in Men with Localised Prostate Cancer. *J Clin Med [Internet]*. 2023;12(7). Available from: <https://www.embase.com/search/results?subaction=viewrecord&id=L2022558771&from=export>
 20. Colen RR, Vangel M, Wang J, Gutman DA, Hwang SN, Wintermark M et al. Imaging genomic mapping of an invasive MRI phenotype predicts patient outcome and metabolic dysfunction: A TCGA glioma phenotype research group project. *BMC Med Genomics [Internet]*. 2014;7(1). Available from: <https://www.embase.com/search/results?subaction=viewrecord&id=L53176027&from=export>
 21. Pan S, Cox H, Willmott J, Mundt E, Gorringer H, Landon M, et al. Discordance between germline genetic findings and abnormal tumor immunohistochemistry staining of mismatch repair proteins in individuals with suspected lynch syndrome. *Front Oncol*. 2023;13:1069467.
 22. Smithgall MC, Remotti H, Hsiao SJ, Mansukhani M, Liu-Jarin X, Fernandes H. Investigation of discrepant mismatch repair immunohistochemistry and microsatellite instability polymerase chain reaction test results for gynecologic cancers using next-generation sequencing. *Hum Pathol*. 2022;119:41–50.
 23. Dix-Peek T, Phakathi BP, van den Berg EJ, Dickens C, Augustine TN, Cubasch H, et al. Discordance between PAM50 intrinsic subtyping and immunohistochemistry in South African women with breast cancer. *Breast Cancer Res Treat*. 2023;199(1):1–12.
 24. Cho N. Breast Cancer Radiogenomics: Association of Enhancement Pattern at DCE MRI with Deregulation of mTOR Pathway. *Radiology*. 2020;296(2):288–9.
 25. Kim HK, Park KH, Kim Y, Park SE, Lee HS, Lim SW, et al. Discordance of the PAM50 intrinsic subtypes compared with immunohistochemistry-based surrogate in breast Cancer patients: potential implication of genomic alterations of discordance. *Cancer Res Treat*. 2019;51(2):737–47.
 26. Hoshino I, Yokota H. Radiogenomics of gastroenterological cancer: the dawn of personalized medicine with artificial intelligence-based image analysis. *Ann Gastroenterol Surg*. 2021;5(4):427–35.
 27. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging—how-to guide and critical reflection. *Insights into Imaging*. 2020;11(1):91.
 28. Zhang W, Guo Y, Jin Q. Radiomics and its feature selection: a review. *Symmetry*. 2023;15(10):1834.
 29. Bera K, Braman N, Gupta A, Velcheti V, Madabhushi A. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat REVIEWS Clin Oncol*. 2022;19(2):132–46.
 30. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. (Aerts HJWL et al. hugo_aerts@dfci.harvard.edu; Velazquez E.R.; Leijenaar R.T.H.; Parmar C.; Cavalho S.; Hoebers F.; Dekker A.; Lambin P.) Department of Radiation Oncology (MAASTRO), Research Institute GROW, Maastricht University, 6229ET Maastricht, Netherlands):4006.
 31. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5(1):4006.
 32. Gevaert O, Xu J, Hoang CD, Leung AN, Xu Y, Quon A, et al. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results. *Radiology*. 2012;264(2):387–96.
 33. Zangfardino M, Pane K, Mirabelli P, Salvatore M, Franzese M. TCGA-TCIA impact on Radiogenomics Cancer Research: a systematic review. *Int J Mol Sci*. 2019;20(23).
 34. Lo Gullo R, Daimiel I, Morris EA, Pinker K. Combining molecular and imaging metrics in cancer: radiogenomics. *Insights Imaging*. 2020;11(1):1.
 35. Woznicki P, Laqua FC, Al-Haj A, Bley T, Baeßler B. Addressing challenges in radiomics research: systematic review and repository of open-access cancer imaging datasets. *Insights into Imaging*. 2023;14(1):216.
 36. Vadapalli S, Abdelhalim H, Zeeshan S, Ahmed Z. Artificial intelligence and machine learning approaches using gene expression and variant data for personalized medicine. *Brief Bioinform*. 2022;23(5):bbac191.
 37. Vial A, Stirling D, Field M, Ros M, Ritz C, Carolan M et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review. *Translational Cancer Research [Internet]*. 2018 Jun [cited 2024 Oct 9];7(3). Available from: <https://tcrgroups.org/article/view/21823>
 38. Wong EY, Chu TN, Ladi-Syedean SS. Genomics and artificial intelligence: prostate cancer. *Urol Clin North Am*. 2024;51(1):27–33.
 39. Unger M, Kather JN. Deep learning in cancer genomics and histopathology. *Genome Med*. 2024;16(1):44.
 40. Ligon K, Lupo J, Molinaro A, Block S, Charbonneau S, Geduldig J, The Ivy glioblastoma patient atlas—a novel clinical and radio-genomics resource for early phase clinical trial design and interpretation. *Neuro-Oncology*. Wen et al. P) Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, United States):vi159.
 41. Jain R. Practical radiogenomics: lessons learnt from TCGA/TCIA and who Glioma classification update. *Neuroradiology*. 2018;60(1):87.
 42. Andreassen CN, Rosenstein BS, Kerns SL, Ostrer H, De Ruysscher D, Cesaretti JA, et al. Individual patient data meta-analysis shows a significant association between the ATM rs1801516 SNP and toxicity after radiotherapy in 5456 breast and prostate cancer patients. *Radiother Oncol*. 2016;121(3):431–9.
 43. Arefan D, Hausler RM, Sumkin JH, Sun M, Wu S. Predicting cell invasion in breast tumor microenvironment from radiological imaging phenotypes. *BMC Cancer*. 2021;21(1):370.
 44. Wang Z, Yuan Y, Cui T, Xu B, Zou Z, Xu Q, et al. Survival and immune microenvironment prediction of glioma based on MRI imaging genomics method: a retrospective observational study. *Neurosurg Rev*. 2025;48(1):18.

45. Guo Y, Li T, Gong B, Hu Y, Wang S, Yang L et al. From images to genes: Radiogenomics based on artificial intelligence to achieve non-invasive precision medicine in cancer patients. *Advanced Science*. n/a(n/a):2408069.
46. Koçak B, Durmaz EŞ, Ateş E, Kılıçkesmez Ö. Radiomics with artificial intelligence: a practical guide for beginners. *Diagn Interv Radiol*. 2019;25(6):485–95.
47. Avanzo M, Wei L, Stancanelli J, Vallières M, Rao A, Morin O, et al. Machine and deep learning methods for radiomics. *Med Phys*. 2020;47(5):e185–202.
48. Cai L, Lambregts DMJ, Beets GL, Mass M, Pooch EHP, Guérendel C, et al. An automated deep learning pipeline for EMVI classification and response prediction of rectal cancer using baseline MRI: a multi-centre study. *npj Precis Onc*. 2024;8(1):1–11.
49. Gasperskaja E, Kučinskas V. The most common technologies and tools for functional genome analysis. *Acta Med Lit*. 2017;24(1):1–11.
50. Swanton C, Bernard E, Abbosh C, André F, Auwerx J, Balmain A, et al. Embracing cancer complexity: Hallmarks of systemic disease. *Cell*. 2024;187(7):1589–616.
51. Figiel S, Yin W, Doultinos D, Erickson A, Poulou N, Singh R, et al. Spatial transcriptomic analysis of virtual prostate biopsy reveals confounding effect of tissue heterogeneity on genomic signatures. *Mol Cancer*. 2023;22(1):162.
52. Davey MG, Richard V, Lowery AJ, Kerin MJ. OncotypeDX® Recurrence score in BRCA mutation carriers: a systematic review and meta-analysis. *Eur J Cancer*. 2021;154:209–16.
53. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–7.
54. Khalili N, Kazerooni AF, Familiar A, Halder D, Kraya A, Foster J, et al. Radiomics for characterization of the glioma immune microenvironment. *npj Precis Onc*. 2023;7(1):1–14.
55. Gu Y, Huang H, Tong Q, Cao M, Ming W, Zhang R, et al. Multi-view radiomics feature fusion reveals distinct immuno-oncological characteristics and clinical prognoses in hepatocellular carcinoma. *Cancers (Basel)*. 2023;15(8):2338.
56. Lee HW, Cho H, ho, Joong JG, Jeon HG, Jeong BC, Jeon SS, et al. Integrative radiogenomics approach for risk assessment of post-operative metastasis in pathological T1 renal cell carcinoma: a pilot retrospective cohort study. *Cancers (Basel)*. 2020;12(4):866.
57. Jamialahmadi H, Khalili-Tanha G, Nazari E, Rezaei-Tavirani M. Artificial intelligence and bioinformatics: a journey from traditional techniques to smart approaches. *Gastroenterol Hepatol Bed Bench*. 2024;17(3):241–52.
58. Yang H, Gou X, Feng C, Zhang Y, Chai F, Hong N, et al. Computed tomography-detected extramural venous invasion-related gene signature: a potential negative biomarker of immune checkpoint inhibitor treatment in patients with gastric cancer. *J Transl Med*. 2023;21(1):4.
59. Yang H, Gou X, Feng C, Zhang Y, Sun B, Peng P, et al. Overall survival prediction of gastric cancer using the gene signature of CT-detected extramural venous invasion combined with M2 macrophages infiltration. *J Transl Med*. 2024;22(1):829.
60. Gao B, Gou X, Feng C, Zhang Y, Gu H, Chai F, et al. Identification of cancer-associated fibroblast subtypes and distinctive role of MFAP5 in CT-detected extramural venous invasion in gastric cancer. *Transl Oncol*. 2025;51:102188.
61. Fan M, Wang K, Zhang Y, Ge Y, Lü Z, Li L. Radiogenomic analysis of cellular tumor-stroma heterogeneity as a prognostic predictor in breast cancer. *J Translational Med*. 2023;21(1):851.
62. Li X, Ding R, Liu Z, Teixeira WMS, Ye J, Tian L et al. A predictive system comprising serum microRNAs and radiomics for residual retroperitoneal masses in metastatic nonseminomatous germ cell tumors. *CR Med [Internet]*. 2024 Dec 17 [cited 2025 Jan 12];5(12). Available from: [https://www.cell.com/cell-report-s-medicine/abstract/S2666-3791\(24\)00614-1](https://www.cell.com/cell-report-s-medicine/abstract/S2666-3791(24)00614-1)
63. Wang S, Zhu C, Jin Y, Yu H, Wu L, Zhang A, et al. A multi-model based on radiogenomics and deep learning techniques associated with histological grade and survival in clear cell renal cell carcinoma. *Insights Imaging*. 2023;14(1):207.
64. Raffy P, Pambrun JF, Kumar A, Dubois D, Patti JW, Cairns RA, et al. Deep learning body region classification of MRI and CT examinations. *J Digit Imaging*. 2023;36(4):1291–301.
65. Sweeney SM, Hamadeh HK, Abrams N, Adam SJ, Brenner S, Connors DE, et al. Challenges to using big data in cancer. *Cancer Res*. 2023;83(8):1175–82.
66. Kocak B, Baessler B, Bakas S, Cuocolo R, Fedorov A, Maier-Hein L, et al. Check-List for Evaluation of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMI. *Insights into Imaging*. 2023;14(1):75.
67. Oliva A, Kaphle A, Reguant R, Sng LMF, Twine NA, Malakar Y, et al. Future-proofing genomic data and consent management: a comprehensive review of technology innovations. *GigaScience*. 2024;13:giae021.
68. Kocak B, Keles A, Kose F, Sendur A. Quality of radiomics research: comprehensive analysis of 1574 unique publications from 89 reviews. *Eur Radiol [Internet]*. 2024 Sep 6 [cited 2024 Oct 8]; Available from: <https://doi.org/10.1007/s00330-024-11057-z>
69. Cobo M, Menéndez Fernández-Miranda P, Bastarrika G, Lloret Iglesias L. Enhancing radiomics and deep learning systems through the standardization of medical imaging workflows. *Sci Data*. 2023;10(1):732.
70. Shur JD, Doran SJ, Kumar S, ap Dafydd D, Downey K, O'Connor JPB, et al. Radiomics in oncology: a practical guide. *Radiographics*. 2021;41(6):1717–32.
71. Masoudi S, Harmon SA, Mehravand S, Walker SM, Raviprakash H, Bagci U, et al. Quick guide on radiology image pre-processing for deep learning applications in prostate cancer research. *J Med Imaging (Bellingham)*. 2021;8(1):010901.
72. Zhou M, Leung A, Echegaray S, Gentles A, Shrager JB, Jensen KC, et al. Non-small Cell Lung Cancer Radiogenomics Map identifies relationships between Molecular and Imaging phenotypes with prognostic implications. *Radiology*. 2018;286(1):307–15.
73. Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K, et al. Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal*. 2022;79:102444.
74. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and guidelines for the interpretation and reporting of sequence variants in Cancer. *J Mol Diagn*. 2017;19(1):4–23.
75. Austin-Tse CA, Jobanputra V, Perry DL, Bick D, Taft RJ, Venner E, et al. Best practices for the interpretation and reporting of clinical whole genome sequencing. *Npj Genom Med*. 2022;7(1):1–13.
76. Ma J, He Y, Li F, Han L, You C, Wang B. Segment anything in medical images. *Nat Commun*. 2024;15(1):654.
77. Diaz-Pinto A, Alle S, Nath V, Tang Y, Ihsani A, Asad M, et al. MONAI label: a framework for AI-assisted interactive labeling of 3D medical images. *Med Image Anal*. 2024;95:103207.
78. Garcea F, Serra A, Lamberti F, Morra L. Data augmentation for medical imaging: a systematic literature review. *Comput Biol Med*. 2023;152:106391.
79. Yao P, Shen S, Xu M, Liu P, Zhang F, Xing J, et al. Single model deep learning on imbalanced small datasets for skin lesion classification. *IEEE Trans Med Imaging*. 2022;41(5):1242–54.
80. Shim SO, Hussain L, Aziz W, Alshdadi AA, Alzahrani A, Omar A. Deep learning convolutional neural network ResNet101 and radiomic features accurately analyzes mpMRI imaging to predict MGMT promoter methylation status with transfer learning approach. Volume 34. *INTERNATIONAL JOURNAL OF IMAGING SYSTEMS AND TECHNOLOGY*; 2024. 2.
81. Jee J, Fong C, Pichotta K, Tran TN, Luthra A, Waters M, et al. Automated real-world data integration improves cancer outcome prediction. *Nature*. 2024;636(8043):728–36.
82. van der Velden BHM, Kuijff HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med Image Anal*. 2022;79:102470.
83. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtman S. Practical guide to SHAP analysis: explaining supervised machine learning model predictions in drug development. *Clin Transl Sci*. 2024;17(11):e70056.
84. Severn C, Suresh K, Görg C, Choi YS, Jain R, Ghosh D. A pipeline for the implementation and visualization of explainable machine learning for medical imaging using radiomics features. *Sensors*. 2022;22(14):5205.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.