Radiomics: a quantitative imaging biomarker in precision oncology

Ashish Kumar Jha\textsuperscript{a,b,c,*}, Sneha Mithuna\textsuperscript{b,c}, Nilendu C. Purandare\textsuperscript{b,c}, Rakesh Kumar\textsuperscript{d,†}, Venkatesh Rangarajanc\textsuperscript{b,†}, Leonard Weea\textsuperscript{a,†} and Andre Dekke\textsuperscript{a,†}

Cancer treatment is heading towards precision medicine driven by genetic and biochemical markers. Various genetic and biochemical markers are utilized to render personalized treatment in cancer. In the last decade, noninvasive imaging biomarkers have also been developed to assist personalized decision support systems in oncology. The imaging biomarkers i.e., radiomics is being researched to develop specific digital phenotype of tumor in cancer. Radiomics is a process to extract high throughput data from medical images by using advanced mathematical and statistical algorithms. The radiomics process involves various steps i.e., image generation, segmentation of region of interest (e.g. a tumor), image preprocessing, radiomic feature extraction, feature analysis and selection and finally prediction model development. Radiomics process explores the heterogeneity, irregularity and size parameters of the tumor to calculate thousands of advanced features. Our study investigates the role of radiomics in precision oncology. Radiomics research has witnessed a rapid growth in the last decade with several studies published that show the potential of radiomics in diagnosis and treatment outcome prediction in oncology. Several radiomics based prediction models have been developed and reported in the literature to predict various prediction endpoints i.e., overall survival, progression-free survival and recurrence in various cancer i.e., brain tumor, head and neck cancer, lung cancer and several other cancer types. Radiomics based digital phenotypes have shown promising results in diagnosis and treatment outcome prediction in oncology. In the coming years, radiomics is going to play a significant role in precision oncology. Nucl Med Commun 43: 483–493 Copyright © 2022 Wolters Kluwer Health, Inc. All rights reserved.

Keywords: artificial intelligence, imaging biomarker, precision oncology, radiomics

*Department of Radiation Oncology (Maastro), GROW School for Oncology, Maastricht University Medical Centre+, The Netherlands, \textsuperscript{b}Department of Nuclear Medicine and Molecular Imaging, Tata Memorial Hospital, \textsuperscript{c}Homi Bhabha National Institute (HBNI), Deemed University, Mumbai and \textsuperscript{d}Department of Nuclear Medicine, All India Institute of Medical Science, New Delhi, India

Correspondence to Ashish Kumar Jha, PGDMRIT, PGDBI, MSc, Department of Radiation Oncology (Maastro), GROW School for Oncology, Maastricht University Medical Centre+, 6229 ET Maastricht, The Netherlands
Tel: +31 9969483165; fax: +31 9969483165; e-mail: a.jha@maastrichtuniversity.nl

*Ashish Kumar Jha is the first author. 
†Rakesh Kumar, V. Rangarajan, L. Wee and Andre Dekker are equal senior authors.

Received 2 October 2021 Accepted 27 December 2021

Introduction
Cancer is caused by genetic mutations leading to uncontrolled growth of tissue and the cells can leave the tissue colony and metastasize in other parts of the body [1]. Conventionally, cancer is treated by surgery or chemotherapy or radiotherapy or combinations of these [2]. Often the selection of treatment depends upon the type of tumor, stage of the disease and general condition of the patient [2]. Although, clinicians consider these factors to decide the course of treatment, on several occasions these treatments fail [3]. This led to the evolution of personalized medicine in oncology [4]. Personalized oncology works on the principle of identification of subgroups of patients in particular disease types [4,5]. Many biomarkers and gene mutations have been investigated to identify the subgroups of the patients in various cancers and targeted drugs for those subgroups [5,6]. For example, by sequencing and in situ hybridization techniques a patient subgroup with epidermal growth factor receptor mutation can be identified in non-small cell lung cancer patients. These high-risk patients do not respond well to conventional treatment options but show good response with targeted therapies like Erlotinib, Gefitinib, Afatinib and similar drugs [6–9]. Precision oncology has the potential to personalize the screening, risk stratifications, treatment selection and response assessment [4,5]. Although most approaches towards precision oncology are centered on biomarkers and genetic mutation assessments [7], artificial intelligence (AI) driven technologies are also being explored to improve the accuracy of precision oncology [10–12]. This technology-driven approach has also been tested in various fields in precision oncology, that is, screening, risk stratifications, treatment selection and response assessment [10,12]. AI-based precision oncology has achieved success as witnessed in published literature in the last few years. Various imaging biomarkers
are being developed and tested for their utility in precision oncology [13–15]. Those imaging biomarkers are of two types i.e., qualitative (e.g. spiculated margin of tumor, vascularity of tumor, position of tumor and contrast enhancement of the tumor, etc.) and quantitative (e.g., hounsfield unit in computed tomography (CT), standardized uptake value and total lesion glycolysis in PET) [16]. In the last few years another kind of imaging biomarker, i.e., radiomic features are extracted from the medical images and being tested in precision oncology [16,17]. The aim of this study is to review the radiomic process and its role in precision oncology and secondary aim was to investigate the growth of radiomics research in the last two decades.

**Radiomics**

Radiomics as a word was first used by Lambin et al. in 2012 in order to describe the quantification of medical imaging data [17]. Radiomics is a process to extract high throughput data from medical images like CT, PET, MRI or SPECT by using advanced mathematical and statistical analysis of images [16,17]. The Radiomics process explores the heterogeneity, irregularity and size parameters of the tumor to calculate thousands of advanced features [16–18]. There are mainly two types of radiomics, i.e., handcrafted radiomics and deep learning-based radiomics. Here in this manuscript mainly we will discuss the first form of radiomics i.e. hand crafted radiomics and we will address these by the term radiomics itself [19].

**Radiomics process**

The Radiomics process involves various steps, i.e., image generation, segmentation of the region of interest (ROI) (e.g. a tumor), image preprocessing, radiomic feature extraction, feature analysis and selection and prediction model development [16–18]. The stepwise radiomic process is shown in Fig. 1.

**Image generation**

Medical equipment like CT, PET, MRI and SPECT are used to image the patient and three-dimensional images are generated by sophisticated reconstruction techniques. These images are archived in image repository, that is, picture archiving communication system (PACS) for future utilization.

**Segmentation**

The images are transferred to the workstations and the ROI is delineated surrounding the tumor, to extract radiomic features from that part of the image. The ROI is generated by medical experts or physicists and typically stored as DICOM RT structure or Segmentation.

**Preprocessing of image**

Image preprocessing involves various steps performed on images and the ROI. As an example, the following steps are typically performed before radiomic extraction from the medical images [20].

**Interpolation**

Medical images are reconstructed and represented in three-dimensional matrices with one unit of the matrix called a voxel. Often voxels are not isotropic and to extract textural radiomic features, the voxels are often resampled or interpolated into isotropic voxels.

**Resegmentation**

The original ROI defined by expert or by automated segmentation is utilized to generate a morphological mask and intensity mask. The morphological mask is the original mask. The intensity mask is resegmented, which contains selected voxel inside or outside the morphological mask.

**Region of interest extraction**

Many features do not require voxels outside the ROI; hence the image volume is extracted for the image based on the ROI of intensity mask.

**Intensity discretization**

Medical images contain noise and often quantization of image intensities is performed to suppress the noise inside the ROI to calculate the texture features. Two
approaches are used for intensity discretization, that is, (1) fixed number of bins and (2) fixed bin width.

**Radiomic feature extraction**

Automatic extraction of radiomic features is performed in this step. Thousands of radiomic features are generated in this step which is further processed in the radiomic analysis step.

**Radiomic analysis and feature selection**

While sometimes 1000+ features are extracted from medical images; these are not all useful for phenotyping a particular disease or for the development of an outcome prediction model. Many features are redundant and many have no association with the particular disease or outcome. Various statistical tests can be performed for feature reduction [21]. Hierarchical clustering, Spearman correlation, Pearson correlation paired t-test are performed to eliminate the redundancy of the feature; forward, backward feature selection, Least Absolute Shrinkage and Selection Operator or recursive feature elimination techniques are used to reduce the dimensionality of the features. Finally, the most appropriate features are selected for disease prognostication or prediction model development for various endpoints like overall survival, recurrence, treatment selection or prediction of treatment outcomes.

**Prediction model development**

Finally, the prediction model is developed and validated by using the selected features. These features may also be combined with clinical features to develop prediction models. Various machine algorithms have been used to develop a prediction model depending upon the need i.e., regression algorithms, Linear and Logistic regression, K-Nearest Neighbor, decision trees algorithms, i.e., Random Forest, Support Vector Machine, Bayesian Network, and deep learning algorithms, i.e., Convolutional Neural Networks, Recurrent Neural Networks and Artificial Neural Networks [22–24]. Radiomic features can be categorized into various groups [18]. Feature groups and the typical number of features extracted using Pyradiomics software [25] are shown in Table 1 (Supplementary material 1, Supplemental digital content 1, http://links.lww.com/NMC/A215).

**Deep learning radiomics workflow**

Recently, an alternative to handcrafted radiomic workflow, a deep learning-based radiomics workflow [26–28] has emerged. A deep learning-based radiomics workflow extracts features from medical images without predefined formulas. Images may be used with or without an ROI for this deep radiomic workflow. Usually, it is a two- or three-step process. Step (1) image data acquisition, (2) segmentation (may or may not be given), (3) development and validation of deep neural networks model. It is not possible in deep learning radiomics to describe features mathematically.

**Radiomics and precision oncology**

Radiomics has witnessed a rapid growth in the last decade with several studies published that show the potential of radiomics in diagnosis and treatment of cancer. Many radiomics based AI decision support systems have been developed in oncology and reported in literature. Figure 2 shows the process of precision oncology leveraging radiomic and artificial intelligence.

In the last few years a new aspect of radiomics, that is, Delta Radiomics is being researched [29]. Delta radiomics comprises extraction and comparison of quantitative features from sequential scans acquired over the course of treatment, which provides information on the efficacy of treatment.

**Methodology**

This study is approved by the Institutional Ethics Committee as a retrospective study. In this study, we have performed literature surveys to find the emerging trend of radiomics based publications in oncology. Our search criteria are optimized to search only those articles, which clearly mention radiomics or related terms like texture analysis in their title. We further extended our search and added year of publication as a criterion to find the total number of publications available on radiomics on PubMed and year-wise distribution of those publications. Furthermore, we added disease and segregated articles based on disease type. To understand the trend of imaging modality used for radiomic study we further included keywords like CT or PET or MRI along with search criteria in all fields. The details of search criteria adopted in this study are mentioned in Tables 2 and 3.

**Results**

We found in total 5243 articles published on radiomics since the year 2000 that satisfied our search criteria. Out of total articles published on radiomics, 624, 2234 and 2110 articles had mention of PET, CT and MRI, respectively (Table 4). The detailed distribution of the publications year wise in all categories are shown in Table 4. There were 123 studies published on radioomic stability study. Maximum 549 articles were published on lung cancer alone followed by 533 articles on GI cancer (Fig. 3).

The percentage of radiomic articles published on CT and MRI are almost the same 45% and 42%, respectively (Fig. 3a). Radiomics articles published on lung and GI cancers contribute approximately 20% of total publications on radiomics (Fig. 3b).

Publication trend on radiomics has shown rapid growth in last decade (Fig. 4a). The trend shows that the yearly publications have increased many folds in the last 5 years (Fig. 4b). A similar growth trend has been...
witnessed in all imaging types (Fig. 4a and b) and all types of cancers (Fig. 5a and b). In our study, we found 85 articles which have utilized all three imaging modalities for radiomic study (Fig. 4c). Figure 4d shows the year wise publication of radiomic articles on stability of radiomic features.

**Discussion**

The utility of radiomic based prediction modeling has been tested widely in diagnosis and treatment of all varieties of solid tumors. Several studies have been performed to differentiate high-grade and low-grade gliomas and to develop various radiomic markers for treatment selection [26,27]. Several studies have shown the association of radiomic features extracted from PET or MRI with survival in glioma [30–39]. Radiomics is widely used in diagnosis and treatment assessment of head and neck cancer [40]. A radiomic signature from PET, MRI and CT has been found to have a significant role in prediction of stage of tumor, HPV status, hypoxia status and gene expression in head and neck cancer [41–51]. Studies have shown the role of radiomics in characterization of sentinel lymph node metastasis in breast cancer noninvasively [52]. The role of radiomics has also been demonstrated by various researchers in breast cancer for response evaluation such as disease-free survival (DFS) [53–57]. The role of radiomics has been widely explored in lung cancer management [58]. Various

<table>
<thead>
<tr>
<th>Type of feature</th>
<th>Feature descriptions</th>
<th>No. of features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shape-based features</strong></td>
<td>Shape features are the descriptors of the three-dimensional size and shape of the ROI and independent from the gray level intensity distribution. These features are only calculated on the original image and mask.</td>
<td>13</td>
</tr>
<tr>
<td><strong>First-order statistics</strong></td>
<td>First-order statistics describe the distribution of voxel intensities within the ROI region of the image.</td>
<td>17</td>
</tr>
<tr>
<td><strong>GLRLM</strong></td>
<td>Gray Level Run Length Matrix (GLRLM) assesses the distribution of discretized gray levels in an image or in a stack of images assesses run lengths.</td>
<td>16</td>
</tr>
<tr>
<td><strong>GLCM</strong></td>
<td>Gray Level Co-Occurrence Matrix (GLCM) expresses how combinations of discretized intensities of neighboring voxels in a 3D volume, are distributed along with one of the image directions.</td>
<td>22</td>
</tr>
<tr>
<td><strong>GLSZM</strong></td>
<td>Gray Level Size Zone Matrix (GLSZM) counts the number of groups/zones of linked voxels with identical discretized gray level.</td>
<td>16</td>
</tr>
<tr>
<td><strong>NGTDM</strong></td>
<td>Neighboring Gray Tone Difference Matrix (NGTDM) contains the sum of gray level differences of voxels with discretized gray level and the average discretized gray level of neighboring voxels within a Chebyshev distance δ.</td>
<td>5</td>
</tr>
<tr>
<td><strong>GLDM</strong></td>
<td>Gray Level Dependence Matrix (GLDM) quantifies gray level dependencies in an image in terms of the number of connected voxels within distance δ that are dependent on the central voxel.</td>
<td>14</td>
</tr>
<tr>
<td><strong>LoG features</strong></td>
<td>A Laplacian of Gaussian (LoG) filter is applied on the original image and one set of derived images is generated for each sigma value specified. Usually, 1-5 sigma values are used, we use 3 sigma values 1, 2, 3 and three sets of derived images are produced. Subsequently, radiomic features are extracted from these image sets.</td>
<td>270</td>
</tr>
<tr>
<td><strong>Wavelet features</strong></td>
<td>Wavelet transformation of image is performed using the three-dimensional wavelet decomposition and 8 sets of images are generated from the original image set. Radiomic features are extracted for transformed image sets.</td>
<td>720</td>
</tr>
</tbody>
</table>

Mechanism to deliver personalized medicine leveraging the machine learning and artificial intelligence to decode the digital signature of the individual patient.
Table 2. The term and search criteria used to select study based on the above-mentioned criteria

<table>
<thead>
<tr>
<th>Modality</th>
<th>Search Criteria</th>
</tr>
</thead>
</table>

Table 3. The term and search criteria used to select radiomic studies published on various cancer types

<table>
<thead>
<tr>
<th>Disease site</th>
<th>Search criteria</th>
</tr>
</thead>
</table>

Table 4. The total and year wise publications on radiomics in oncology

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1549</td>
<td>1277</td>
<td>798</td>
<td>535</td>
<td>298</td>
<td>175</td>
<td>132</td>
<td>5243</td>
</tr>
<tr>
<td>CT</td>
<td>733</td>
<td>592</td>
<td>343</td>
<td>236</td>
<td>132</td>
<td>70</td>
<td>38</td>
<td>2234</td>
</tr>
<tr>
<td>PET</td>
<td>169</td>
<td>151</td>
<td>105</td>
<td>71</td>
<td>43</td>
<td>31</td>
<td>14</td>
<td>624</td>
</tr>
<tr>
<td>MRI</td>
<td>645</td>
<td>527</td>
<td>345</td>
<td>205</td>
<td>108</td>
<td>74</td>
<td>43</td>
<td>2110</td>
</tr>
<tr>
<td>CT-PET</td>
<td>142</td>
<td>118</td>
<td>82</td>
<td>46</td>
<td>33</td>
<td>26</td>
<td>5</td>
<td>475</td>
</tr>
<tr>
<td>PET-MR</td>
<td>42</td>
<td>33</td>
<td>21</td>
<td>17</td>
<td>9</td>
<td>6</td>
<td>2</td>
<td>135</td>
</tr>
<tr>
<td>CT-MRI</td>
<td>91</td>
<td>58</td>
<td>42</td>
<td>15</td>
<td>11</td>
<td>8</td>
<td>2</td>
<td>234</td>
</tr>
<tr>
<td>CT-MRI-PET</td>
<td>32</td>
<td>19</td>
<td>16</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>85</td>
</tr>
<tr>
<td>Stability</td>
<td>47</td>
<td>26</td>
<td>29</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>123</td>
</tr>
<tr>
<td>Head &amp; neck cancer</td>
<td>80</td>
<td>74</td>
<td>45</td>
<td>36</td>
<td>24</td>
<td>14</td>
<td>8</td>
<td>307</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>26</td>
<td>31</td>
<td>30</td>
<td>12</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>122</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>155</td>
<td>157</td>
<td>78</td>
<td>71</td>
<td>39</td>
<td>20</td>
<td>10</td>
<td>549</td>
</tr>
<tr>
<td>GI cancer</td>
<td>110</td>
<td>93</td>
<td>58</td>
<td>31</td>
<td>25</td>
<td>11</td>
<td>13</td>
<td>369</td>
</tr>
<tr>
<td>Cervical cancer</td>
<td>182</td>
<td>156</td>
<td>85</td>
<td>51</td>
<td>17</td>
<td>10</td>
<td>11</td>
<td>533</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>38</td>
<td>29</td>
<td>15</td>
<td>12</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>104</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>57</td>
<td>42</td>
<td>38</td>
<td>20</td>
<td>8</td>
<td>10</td>
<td>3</td>
<td>251</td>
</tr>
</tbody>
</table>

studies have been performed to differentiate between benign and malignant tumor, pathology types (i.e. adenocarcinoma or squamous cell carcinoma), EGFR mutation status and various TNM stages [59–65]. Literature published in the last one decade also suggests an increasing role of radiomic features in the prediction of OS, PFS,
Figure 3.

Publications on radiomics: (a) Imaging modality wise distribution of articles; (b) disease wise distribution of articles.

Figure 4.

Figure shows (a) the trend of number of publications on PET, CT and MRI radiomics in oncology over the last two decades, (b) of number of publications on PET, CT and MRI radiomics in oncology since 2015. (c) Venn diagram shows the PET, CT and MRI imaging modality used for radiomic studies; (d) shows the trend of number of published radiomic stability issues since 2015.
DFS, LRR, treatment response, toxicity and quality of life [66–72]. Radiomic features have been explored for the management of colorectal cancer. Various studies have demonstrated the role of radiomic features in the detection of lymph node metastasis, prediction of KRAS/NRAS/ BRAF mutation [73,74]. The role of radiomic features has also been investigated for treatment selection, treatment modification and DFS prediction [73,75–77]. Radiomics has been investigated in prostate cancer management and features extracted from MRI and PET have shown promising results. Several studies have shown the utility of radiomic features in the differentiation between benign and malignant tumor, aggressiveness of tumors and the Gleason Score [78–81]. Many researchers have also shown the utility of radiomic features extracted from MRI and PET to predict biochemical recurrence, PFS and OS [82–84]. GI and liver cancer is another area where the role of radiomics has been investigated in disease management. The role of radiomics has been successfully demonstrated in microvascular invasion detection of liver cancer and differentiation in various kinds of GI malignancies, histology type and TNM staging in the GI cancer [85–91]. Various studies have demonstrated the role of radiomic features in detection of lymph node metastasis, OS, PSF and toxicity prediction in cervical cancer [92–95].

Our study shows an increasing trend of radiomics in oncology in the last decade. The last 5 years witnessed the tremendous growth of radiomic studies in oncology. In all major disease types growth of radiomic studies have been witnessed. Several articles have been published on radiomic stability problems that show the researchers have identified it as a major issue in radiomic implementation.

**Implementation or radiomics based workflow in clinic**

The future of radiomics lies in the clinical application of radiomics. A self-learning model may be developed and implemented in the clinic for participation in the decision support system. There will be requirements of a super-specialized model to address the specific clinical questions. As suggested by Lambin et al., the image archival system, that is, PACS has to be modified to picture archiving and radiomics knowledge systems to store radiomic signatures [16]. The future implementation of the radiomic process may look like Fig. 6.

**Limitations of radiomic implementation**

The main problem of radiomics is its limited repeatability and reproducibility which is thought to be mainly caused by the difference in scanners from different vendors, different acquisition protocols and intra scanner
variations. In our earlier repeatability and reproducibility study, we found that only 10% of CT radiomic features had a good repeatability and reproducibility in a clinical cohort and in phantoms [96]. Traverso et al. in a systematic literature review have also concluded that there are stability issues with majority of radiomic features [97]. In order to harmonize radiomic extraction tools, features and imaging standards, several initiatives are started by various agencies, like The Quantitative Imaging Network (QIN) [98], the Quantitative Imaging Biomarkers Alliance (QIBA) [99], and Quantitative Imaging in Cancer: Connecting Cellular Processes with Therapy (QuIC-ConCePT) [100]. These initiatives are working to standardize imaging and imaging biomarkers. The Image Biomarker Standardization Initiative (IBSI) is another consortium that works towards the harmonization of radiomic features across the globe by minimizing the deviation in imaging and standardizing the radiomic extraction process [101,102]. The radiomics quality score (RQS) is another such initiative proposed by Lambin et al. to address the issues related to radiomic study reporting [16]. Most of these initiatives will assist in advancing the standardization process of imaging biomarkers and are thus expected to address the repeatability and reproducibility challenges currently present in Radiomics.

**Conclusion**

This literature review is suggestive of the increasing role of radiomics in precision oncology. Publications on radiomics have increased many folds in the last 5 years. Initiatives like QIN, QIBA, QuIC-ConCePT, IBSI and RQS will be able to address repeatability and reproducibility of radiomic features. We envision that radiomics is going to play a pivotal role in phenotyping the cancer and guide cancer management to provide more precise treatments to patients in a true clinical environment soon.

**Acknowledgements**

This work was supported by Indo-Dutch NWO research grant BIONIC (grant no.: 629.002.205, A.D.) and Ministry of Electronics and Information Technology (MietY) (IN) (grant no.: 13(2)-2015-CC-BT; V.R.).

Ethics approval: Approved by Institutional Ethics Committee with waiver of informed consent.
Consent for publication: All the authors consented to the publication.

Conflicts of interest
There are no conflicts of interest.

References
computed tomography and [18F]-Fluorodeoxyglucose positron emission tomography radiomics features. *Radiother Oncol* 2018; 127:36–42.


98 National Cancer Institute, Division of Cancer Treatment & Diagnosis. Quantitative Imaging Network (QIN) [online], https://imaging.cancer.gov/programs_resources/specialized_initiatives/qin.htm. last accessed on 15-05-2021


100 QUantitative Imaging in Cancer: CONnecting CEllular Processes with Therapy, https://cordis.europa.eu/project/id/115151 last accessed on 15-05-2021
