




# Incorporating radiomics into clinical trials: expert consensus endorsed by the European Society of Radiology on considerations for data-driven compared to biologically driven quantitative biomarkers

Laure Fournier<sup>1,2,3</sup> · Lena Costaridou<sup>2,4</sup> · Luc Bidaut<sup>3,5</sup> · Nicolas Michoux<sup>3,6</sup> · Frederic E. Lecouvet<sup>3,6</sup> · Lioe-Fee de Geus-Oei<sup>3,7,8</sup> · Ronald Boellaard<sup>2,9,10</sup> · Daniela E. Oprea-Lager<sup>3,9</sup> · Nancy A Obuchowski<sup>10,11</sup> · Anna Caroli<sup>2,12</sup> · Wolfgang G. Kunz<sup>3,13</sup> · Edwin H. Oei<sup>2,14</sup> · James P. B. O'Connor<sup>2,15</sup> · Marius E. Mayerhoefer<sup>2,16</sup> · Manuela Franca<sup>2,17</sup> · Angel Alberich-Bayarri<sup>2,18</sup> · Christophe M. Deroose<sup>3,19,20</sup> · Christian Loewe<sup>2,21</sup> · Rashindra Manniesing<sup>2,22</sup> · Caroline Caramella<sup>3,23</sup> · Egesta Lopci<sup>3,24</sup> · Nathalie Lassau<sup>2,3,10,25</sup> · Anders Persson<sup>2,26</sup> · Rik Achten<sup>2,27</sup> · Karen Rosendahl<sup>2,28</sup> · Olivier Clement<sup>1,2</sup> · Elmar Kotter<sup>2,29</sup> · Xavier Golay<sup>2,10,30</sup> · Marion Smits<sup>2,3,14</sup> · Marc Dewey<sup>2,31</sup> · Daniel C. Sullivan<sup>2,10,32</sup> · Aad van der Lugt<sup>2,14</sup> · Nandita M. deSouza<sup>2,3,10,33</sup>  · European Society of Radiology<sup>34</sup>

Received: 14 July 2020 / Revised: 16 November 2020 / Accepted: 3 December 2020

© The Author(s) 2021, corrected publication 2021

## Abstract

Existing quantitative imaging biomarkers (QIBs) are associated with known biological tissue characteristics and follow a well-understood path of technical, biological and clinical validation before incorporation into clinical trials. In radiomics, novel data-driven processes extract numerous visually imperceptible statistical features from the imaging data with no a priori assumptions on their correlation with biological processes. The selection of relevant features (radiomic signature) and incorporation into clinical trials therefore requires additional considerations to ensure meaningful imaging endpoints. Also, the number of radiomic features tested means that power calculations would result in sample sizes impossible to achieve within clinical trials. This article examines how the process of standardising and validating data-driven imaging biomarkers differs from those based on biological associations. Radiomic signatures are best developed initially on datasets that represent diversity of acquisition protocols as well as diversity of disease and of normal findings, rather than within clinical trials with standardised and optimised protocols as this would risk the selection of radiomic features being linked to the imaging process rather than the pathology. Normalisation through discretisation and feature harmonisation are essential pre-processing steps. Biological correlation may be performed after the technical and clinical validity of a radiomic signature is established, but is not mandatory. Feature selection may be part of discovery within a radiomics-specific trial or represent exploratory endpoints within an established trial; a previously validated radiomic signature may even be used as a primary/secondary endpoint, particularly if associations are demonstrated with specific biological processes and pathways being targeted within clinical trials.

## Key Points

- *Data-driven processes like radiomics risk false discoveries due to high-dimensionality of the dataset compared to sample size, making adequate diversity of the data, cross-validation and external validation essential to mitigate the risks of spurious associations and overfitting.*
- *Use of radiomic signatures within clinical trials requires multistep standardisation of image acquisition, image analysis and data mining processes.*
- *Biological correlation may be established after clinical validation but is not mandatory.*

**Keywords** Radiology · Statistics and numerical data · Standardization · Validation studies · Clinical trial

✉ Nandita M. deSouza  
nandita.deSouza@icr.ac.uk

Extended author information available on the last page of the article

## Abbreviations

ADC	Apparent diffusion coefficient
CE	Conformite Europeenne

CNN	Convolutional neural networks
CT	Computerised tomography
DL	Deep learning
DR	Deep radiomics
EGFR	Epidermal growth factor receptor
FDG	Fluorodeoxyglucose
IBSI	Image biomarker standardisation initiative
MeSH	Medical Subject Headings
MRI	Magnetic resonance imaging
PET	Positron emission tomography
QA	Quality assurance
QC	Quality control
QIBs	Quantitative imaging biomarkers
SPECT	Single photon emission computed tomography
SUV	Standardised uptake value
VOI	Volume of interest

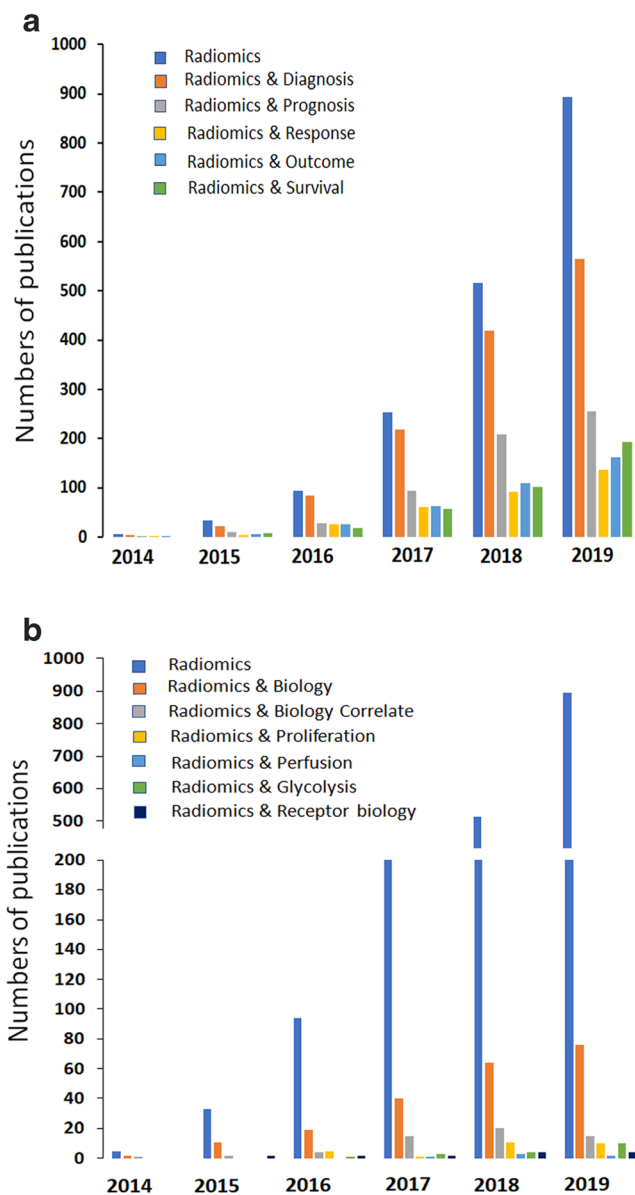
## Introduction

Quantitative imaging biomarkers (QIBs) are associated with tissue characteristics that are altered by disease and its treatment. Necrosis decreases tissue cellularity and increases water content manifesting as an increase in T2 [1], a reduction in glucose uptake [2] and an increase in elasticity [3]. Perfusion imaging detects and characterises hypervascular lesions such as cancers, or monitors the effect of anti-angiogenic drugs [4, 5]. Implementation of QIBs into clinical trials follows a well-defined path from discovery, through a process of technical and biological validation, to implementation and clinical validation. A roadmap defining the process was published as a consensus statement from multiple stakeholders [6]. Despite this, QIBs have been slow to be adopted as trial endpoints because of the relative complexity of imaging protocols and variability of the quantified output under differing conditions (e.g. hardware, software, protocol and observer variability) [7].

Recently, a new approach to derive imaging biomarkers has been advocated through the concept of radiomics [8, 9]. This data-driven framework ‘discovers’ quantitative information within images by extracting high-dimensional data (‘features’) beyond that visually perceptible, using computational statistics (often based on machine learning algorithms) to predict or establish association with a meaningful clinical endpoint [10, 11]. Technical and clinical performance of the ‘radiomic signature’ (specific combination of mathematically derived features) determines its appropriateness. If considered necessary, a link to a biological process is explored a posteriori [12]. Radiomic signatures have been associated with outcome or response [13], and may be used together with clinical, histological and genomic metrics as part of a nomogram of features [14]. The exponential rise in publications involving data-driven biomarkers has not been accompanied

by a mechanism-based understanding of their nature but focuses on their ability to classify disease and patient outcome (Fig. 1). Radiomics has been used for detecting cancer [15], cancer staging [16], performing classifications [17], assessing response to chemotherapy [18], radiation therapies [19–22], immunotherapy [23–26] and predicting/prognosing survival [27].

A major disadvantage of a non-mechanistic data-driven approach is that random chance associations may occur. Most studies look at the associations between a large number of features extracted from discretised images and prognosis/



**Fig. 1** Increase in radiomics related publications over last 6 years (a) by patient status/outcome and (b) by biological association using data extracted from PubMed using the indicated MeSH terms. The exponential increase in radiomics publications relates mainly to usage as indicated in a, and not to their underlying biological associations as indicated in b

response/outcome in an inadequate number of samples. For biomarker profiles that rely on statistical rather than biological associations, generalisation and scalability to multicentre trials requires more than a simple standardisation process. Also, their validation pathway needs to incorporate measures that may differ substantially from traditionally accepted methods. This article prepared by imaging experts from the European Society of Radiology EIBALL (European Imaging Biomarker Alliance) and the EORTC (European Organisation for Research and Treatment of Cancer) Imaging Group with representatives from QIBA (Quantitative Imaging Biomarkers Alliance) examines how the process of standardising and validating data-driven imaging biomarkers differs from those based on biological associations, and what measures need to be considered when implementing them into clinical trials and, eventually, into clinical routine. Structured discussions were conducted via teleconferencing and written communications.

## Standardising the radiomics process for clinical trials

Radiomics analyses rely on image acquisition, image analysis and computational statistics [28], so standardisation of these domains is mandatory prior to their validation (Table 1). As radiomics analyses have been applied to CT [29–31], MRI

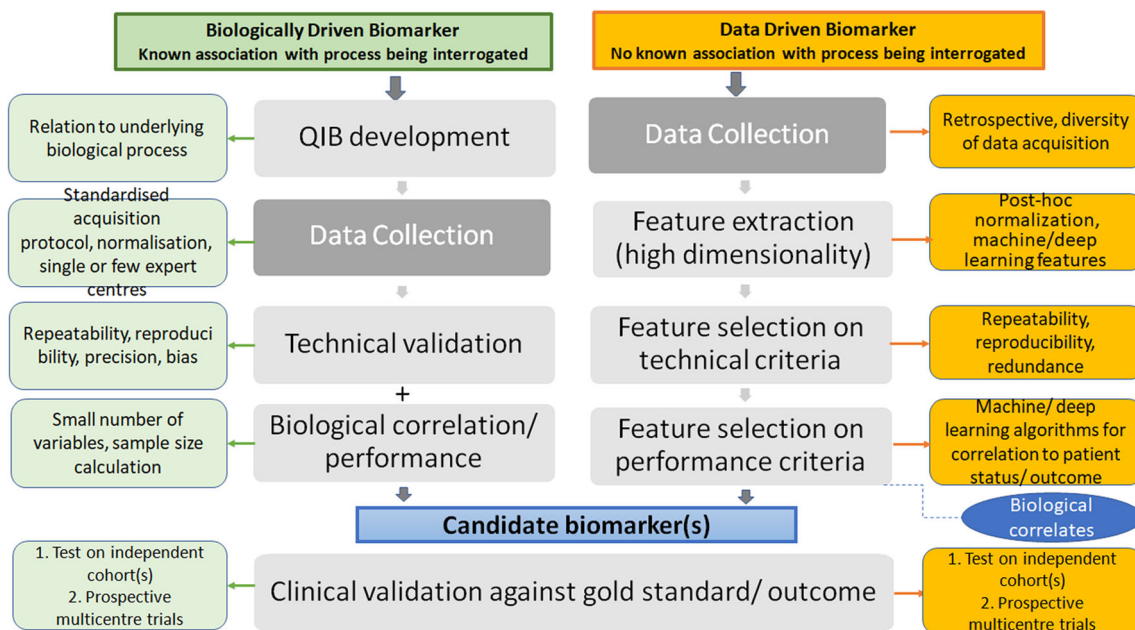
[32–36], nuclear medicine using FDG-PET [37–42] and other tracers [43, 44], and ultrasound [45], image acquisition standardisation needs to consider modality, scanner and scan protocol. Standardisation of image analysis needs to consider software (consistency of technical implementation) and subjectivity (human interaction). Standardisation of computational statistics needs to consider adequacy, performance and requirements for validation of algorithms and models (Fig. 2).

**Image acquisition and normalisation** An element of diversity of acquisition protocols or machines is advantageous at the discovery phase of data-driven biomarkers so that the identified radiomic signatures used in clinical trials are robust enough across a range of platforms [46]. Datasets utilised for radiomic signature development must be representative of the disease and capture the variability and severity for which they will be used. Within a clinical trials framework, as with previously published recommendations and guidelines [6, 47–49], an optimised tightly controlled standardised imaging protocol ensures image quality (low level of noise, artifact-free, spatial resolution) and stability over time, with known intra- and inter-site reproducibility that does not exceed the expected level of change associated with the trial intervention [50]. Phantom studies are limited for quality control of high-dimensionality information [51] because a suitable phantom would need to exhibit high-dimensionality in a realistic setting and cover the requirements of each type of feature.

**Table 1** Comparison of standardisation steps for biologically driven and data-driven biomarkers (QA, quality assurance; QC, quality Control; VOI, volume of interest)

Steps	Biologically driven quantitative biomarkers	Data-driven quantitative biomarkers
Image acquisition	<ul style="list-style-type: none"> <li>Standardised protocols (single and multicentre)</li> <li>QA/QC process across instruments, sites</li> <li>Stability of measurement monitored with phantom studies; may be strengthened by human subject test-retest</li> </ul>	<ul style="list-style-type: none"> <li>Non-standardised protocols in discovery phase followed by standardised protocols within trials</li> <li>QA/QC process across instruments, sites</li> <li>Stability of measurement requires human subject test-retest</li> </ul>
VOI delineation	<ul style="list-style-type: none"> <li>Can be manual or semi-automated</li> <li>Can be machine-learned</li> <li>Deep learning available but infrequently used</li> </ul>	<ul style="list-style-type: none"> <li>Can be manual or semi-automated</li> <li>Can be machine-learned</li> <li>Can be derived from fully convolutional neural networks</li> </ul>
Data analysis	<ul style="list-style-type: none"> <li>Commercial or academic software applicable to datasets regardless of their source</li> </ul>	<ul style="list-style-type: none"> <li>Algorithms used are specific to image datasets and may require adaptation and standardisation for individual situations or new datasets*</li> </ul>
Biomarker extraction	<ul style="list-style-type: none"> <li>Follows standard formula that describes the biological feature (e.g. tissue density, perfusion, diffusion, standardised uptake of radiotracers related to a biological process/receptor status)</li> </ul>	<ul style="list-style-type: none"> <li>Algorithm-based mathematical feature extraction not directly linked to a biological process, followed by selection of feature combination that best separate disease from no disease, good from poor outcome (e.g. shape features such as diameter, sphericity; histogram-derived features such as median, skewness, entropy; texture features such as contrast, homogeneity, Haralick variance)</li> </ul>
Biomarker interpretation	<ul style="list-style-type: none"> <li>Directly linked to biological process</li> </ul>	<ul style="list-style-type: none"> <li>Indirect associations with biological process assumed</li> </ul>

\*<https://ibsi.readthedocs.io/en/latest/>; \*<https://www.lifexsoft.org/>; \*<http://www.eletel.p.lodz.pl/programy/mazda/>; \*<https://nmmitools.org/2019/01/01/pyradiomics/>



**Fig. 2** Pathways comparing processes required for biologically driven and data-driven biomarkers. Biologically driven biomarkers derived from known associations with a specific biological process require a specific predetermined acquisition protocol and image processing technique and involve technical, biological and clinical validation steps with recognised requirements (green boxes). Data-driven biomarkers assume that the statistical features that relate to the biological process or outcome are unknown so that all possible features are extracted from the

images and steps to determine their technical and clinical performance are needed (orange boxes). Feature extraction and selection depend on the data mining process (machine and deep learning algorithms). A training dataset and validation dataset allow selection of most promising feature(s), and an independent test dataset allows evaluation of performance of imaging biomarker. Biological links are explored a posteriori

Basic methods of image normalisation include pixel size resampling by filtering [52] and/or resampling (rescaling) values with respect to global or local mean and standard deviation of reference image/tissue, or by adjusting the histograms [53]. Normalisation methods affect reproducibility of image features [54, 55]. For second-order statistics features, reduction of matrix dimension post-normalisation is needed. This is achieved by discretisation (quantisation, grey-level resampling, histogram re-binning) and reduces noise from clustered intensity values. Choice of the absolute (fixed bin size) or the relative (fixed bin number) method significantly affects the values of texture features and requires optimisation depending on the clinical task at hand [56–58]. Shape features (area, centroid, perimeter, roundness, Feret’s diameter) are less sensitive to differences in intensity values. Both types of features remain dependent on the spatial resolution of the image. Numerical harmonisation of features as an alternative to standardisation of image acquisition and pre-processing is based on transformation of variable feature distributions to a common batch-effect free reference space, to deal with varying imaging conditions [59, 60]

The Image Biomarker Standardization Initiative (IBSI) [61] offers a common reference of definitions and benchmarking of radiomic features and provides recommendations for comprehensive reporting of image acquisition parameters and pre-processing methods.

**Image analysis—segmentation** As with biologically driven biomarkers, manual region of interest delineation introduces inter- and intra-observer variability because of variation in border perception. Observer training and working to protocol assists in this regard. Semi-automated segmentation methods, e.g. region-growing or level set active contour models [62] and deep learning methods [63], are more reproducible [64], but they are dependent on their training set, which may introduce other errors. Quantitative verification metrics [65], such as Dice coefficient, and Hausdorff distance metrics, help determine segmentation reproducibility. Images that require alignment for different time series data, parametric maps and modalities should evaluate deviations in locations (distance) of pairs of homologous landmark points, especially important for non-rigid image registration [66, 67].

**Image analysis—feature extraction** ‘Hand-crafted’ radiomics extracts predefined human-engineered features from the volume-of-interest (VOI) [17]. These include shape characteristics, intensity histogram metrics and texture parameters (local binary patterns, grey-level co-occurrence, run-length, zone-length and neighbourhood different matrices, auto-regressive model, Markov random fields, Riesz wavelets, S-transform, fractals) which require specific assumptions in their computation, so that software

implementations on different platforms (even if all are IBSI compliant) and between different versions of the same software can lead to different results [68]. Recommendations on calculating and reporting radiomic features have been proposed, and both mathematical equations and pre-processing applied should be reported. The information and framework provided through IBSI [61] should also be followed as much as possible to ensure the quality and relevance of the post-processing (denoising, resampling, enhancement, spatial alignment correction, segmentation and feature extraction). Other descriptive (radiologist-scored), functional (SUV, ADC,  $K^{\text{trans}}$ ) or clinical parameters may be added to the radiomic signature if pertinent.

**Computational statistics—feature selection** Several tools are described [69–72]. To identify relevant, non-redundant and stable features with which to build models, three categories of technique are employed. Filter methods (ANOVA, correlation, RELIEF [73]) rely on a criterion function, have low computational cost and are less prone to overfitting, by separating selection from model building; however, they are more unstable to different datasets. Wrapper methods (forward selection, backward elimination, stepwise selection) incorporate a specific machine learning algorithm to eliminate features but have increased computational cost and high probability of overfitting, since model training uses feature combinations that include common features. Embedded methods (LASSO, RIDGE regression) embed features successively and penalise the coefficients of a model that contribute to overfitting at each iteration. They represent a trade-off between filter and wrapper methods.

**Computational statistics—classifier/model** After dimension reduction, selected features are investigated for their association with clinical outcome using tools such as univariable or multivariable logistic regression, decision tree, random forest, support vector machine, neural networks, all described extensively in previous publications [65–68] and used for QIBs and radiomic analyses [24]. Classifiers are differentiated depending on the nature of the clinical outcome, i.e. discrete (mainly binary) or continuous [74, 75]. No tool has proved universally superior and most require a compromise between complexity of tuning versus interpretability of results.

**Computational statistics—deep radiomics (DR)** A recent evolution has been the integration of radiomics with deep learning (DL) [76–78]. ‘Discovery Radiomics’ automatically extracts deep features relevant to a given query (e.g. diagnosis, prognosis) from the data, and the resulting trained model can be applied to complete datasets, avoiding the error-prone segmentation step. As DL can include multiple data types, relevant information in electronic patient records can be exploited.

## Validating the radiomics output

**Technical validation** Following identification of a radiomics signature associated with disease/outcome, two fully independent datasets are needed, one for training and cross-validation (internal validation), and at least one other to test the final model and confirm generalisability and performance (external validation). Both training and testing datasets should be of sufficient uniform quality (data balancing) and representative for the patient population for which the radiomics model is intended. An adequate sample (size and diversity) is essential for the training and validation datasets, with respect to the number and type of features (‘signature’) considered. Testing the model with a dataset containing a different prevalence of cases and/or a high degree of imbalance may result in overoptimistic conclusions. Feature selection avoids overparameterised models, reduces dimensionality of the feature space (data dimension reduction) and ensures that only a small and stable subset of original features relevant to the task are retained. A strategy to cross-validate the structure of the model requires careful considerations regarding sample size, accuracy estimation and the choice of the validation method (hold-out, k-fold cross-validation, bootstrap). Grid searches pose the danger of overfitting, leading to overoptimistic model performance that is not reproduced on other datasets or in clinical practice. Finally, repeatability and reproducibility of the signature in a multicentre context (affected by imaging apparatus, acquisition protocols and analysis methods) is a crucial step in technical validation [79–81]. As with QIBs, radiomics models should be tested with cross-institutional clinical training and testing datasets to guarantee generalisability to representative patient populations.

**Biological validation** Biological correlation with liquid/tissue biopsies may be performed after the technical and clinical validity of a radiomic signature is established but is not mandatory. A radiomic signature that is related to survival outcomes may potentially reflect a tissue phenotype associated with a specific biology. Biological validation reduces the likelihood that radiomic features are selected by statistical chance or may be attributed to the nature of the data sample used for model development. It also offers the opportunity to reduce the number of selected features.

**Clinical validation** The process by which the clinical utility of a single quantitative feature, or multiple features embedded in a statistical model is demonstrated, allowing improvement of health outcomes (improved diagnosis or therapeutic management of a disease or individual patient) is being addressed slowly for radiomics. Following initial ‘discovery’, new and independent datasets are required to replicate the performance of the identified model and validate it clinically. Performance metrics, e.g. sensitivity and specificity, should be evaluated

ideally in prospective trials, or prospectively in the clinic using routinely obtained clinical data (real-life conditions) in order to avoid bias. Table 2 lists some exemplar studies and their clinical use. Broadly speaking, standard recommendations for clinical validation and clinical utility assessment of any QIB should be followed and applied.

## Biological correlates of radiomic features

Images provide an averaged macroscopic view (with large partial volume effects, both in space and time) of the geometry and/or function of the tissue. Radiomic features are statistical descriptors characterising the macroscopic visual aspect of images and only indirectly relate to the microscopic histological characteristics of the imaged tissue. Such features are then used as a statistical/phenomenological description of the outcome, and not embedded into an actual biological/physical model of this outcome that would unambiguously establish causality between features and outcome.

Radiomic information on visually imperceptible phenotypic characteristics such as intensity, shape, size and texture distinguish benign and malignant tumours, likely

reflecting different cellular morphology [101]. In cervix cancer, radiomic features of low-volume tumours with radiomic profiles similar to high-volume tumours had a worse prognosis implying a more aggressive phenotype at an earlier stage [36]. In a lung cancer study, texture entropy and cluster features, as well as voxel intensity variance features, were associated with the immune system, the p53 pathway, pathways involved in cell cycle regulation [102] and for predicting EGFR mutation status [103]. Nevertheless, why specific features are associated with specific pathways remains unexplored and the relationship between radiomic signature and cell morphology, density, distribution pattern, alignment and organelle composition need further elucidation.

Although it is possible to extract mathematically hundreds or thousands of radiomic features from digital images, most studies to date suggest that less than 20 are indicative of unfavourable biology, and these largely relate to shape and textural uniformity. 2D shape features indicate more rapidly progressive disease with reduced overall survival in glioblastoma multiforme [104]. Shape and textural features from CT scans of lung cancer have been shown to predict unfavourable biology (nodal and distant metastases respectively) [105]. In

**Table 2** Exemplar radiomics signature studies and their clinical use

Radiomic analysis	Radiomic feature (process)	Modality	Tissue types investigated	Decision-making role
Second-order statistics	Textural (Haralick, Gabor)	CT [29–31] MRI [24–26] PET/CT [37–42]	Lung, breast, brain, liver, prostate, head and neck, lymph node, cervix	<ul style="list-style-type: none"> <li>• Prognostic</li> <li>• Predictive</li> <li>• Response</li> <li>• Survival</li> <li>• EGFR expression</li> <li>• p53 mutation status</li> </ul>
Higher-order statistics Filter grids extract repetitive or non-repetitive patterns	Wavelets  Laplacian transforms (bandpass filters) Minkowski functions (patterns of voxels with intensity above threshold)  Fractal dimensions (patterns imposed on image and number of grid elements containing voxels of a specified value is computed)	CT [82–87] MRI [88–90] PET/CT [91, 92]  CT [93, 94] MRI [95–97] PET/CT [92]	Lung, oesophagus, brain, pancreas, breast, head and neck  Brain, lung, rectum, cervix, kidney	<ul style="list-style-type: none"> <li>• Diagnostic</li> <li>• Prognostic</li> <li>• Predictive</li> <li>• Response</li> <li>• Survival</li> <li>• Surgical resection margins</li> </ul> Prognostic Response
Delta radiomics	Change in radiomic features	PET/CT [98, 99]	Lung	Response
Dynamic radiomic studies	Pharmacokinetic radiomic features	PET/CT [100]	Lung	Response, data highly correlated to data from static studies

prostate cancer, Gabor textural features (defining spatial frequency patterns within the image) were predictive of Gleason grade on MRI. As gland lumen shape features relate to Gleason grade, discriminability of Gabor features is a likely consequence of variations in gland shape and morphology at the tissue level [106]. In future, prospective selection of a handful of relevant features should become possible to interrogate specific biological processes and pathways being manipulated within clinical trials so that it may be possible for the clinical question to drive the choice of biomarker usage and analysis. However, understanding the biological basis for a biomarker to facilitate its acceptance into clinical practice is not the primary objective of a data-driven process such as radiomics. It may well be that reliable modelling of the outcome with a relatively high and clinically acceptable performance means that biological validation would not be a primary concern [107].

### Limitations of data-driven processes

When defining training datasets for radiomic feature extraction and selection in clinical trials, case-control data may be considered but may underrepresent the disease. Enrichment of training datasets with normal and abnormal cases of varying disease severity is mandatory to achieve appropriate balance. Bias in the training datasets limits generalisability. For example, a radiomic signature developed on lung nodules detected on chest x-rays in a population with a high prevalence of tuberculosis and few cancers will overdiagnose tuberculosis in a population with a high prevalence of cancer. Image acquisition bias (cases recognised as disease acquired with a specific protocol or device) where selected features are linked to image acquisition rather than to image

content may fail to predict disease when applied to an independent population. Manual VOI segmentation and use of locally developed methodology risks discovery of features that are not generalisable and may be influenced by hardware or software-related factors rather than the disease itself. Diverse but balanced image acquisition conditions in the training dataset should counteract these effects. Though balance and diversity are necessary at the discovery stage, it is crucial to evaluate performance only on populations representative of the natural prevalence.

The radiomic process, which tests combinations of hundreds and thousands of parameters, risks false discovery. Traditional statistical corrections for multiple tests would lead to  $p$  values impossible to reach. Strategies to reduce spurious correlations and overfitting include artificially increasing the number of samples by data augmentation (datasets flipped, rotated and deformed to simulate new patients). Cross-validation or bootstrapping are alternative strategies, but an independent dataset to confirm the findings is always required.

### Implementation of radiomics in clinical trials

Although the discovery phase requires image acquisition diversity, standardised protocols, pre- and post-processing methods, tools and algorithms for feature extraction are needed for incorporating into clinical trials and facilitated by centralised data analyses and publicly available analysis software (Table 3). To incorporate radiomics in clinical trials, three potential scenarios can be considered. *Firstly*, where radiomic signature discovery is the objective, a trial should follow the steps described and illustrated (Fig. 2). *Secondly*, a radiomic ‘exploratory end-point’ may form an ancillary

**Table 3** Recommended process for inclusion of data-driven biomarkers into clinical trials

Step	Recommended process for clinical trial inclusion
Image acquisition	Standardised protocol agreed with site with vendor-specific amendments (incl. software version control) to achieve reproducibility of other QIBs within accepted published standards
Image acquisition—normalisation	Raw data saved. Image normalisation predefined
Image analysis—segmentation	If manual or semi-automated, done by centralised/core laboratory by > 1 observer to establish reproducibility. If automated, can be done with CE-marked software with established limits of agreement at local sites
Image analysis—feature extraction	Use of validated features with established error margins, adapted for individual situations. Discard redundant features. Test reproducibility, repeatability within trial setting
Computational statistics—feature and model selection	Based on performance by association with trial endpoint (e.g. response/survival)
Validation	Adequate sample size, test data on samples with similar characteristics, cross-validation strategies, avoid over-fitted models
Biomarker interpretation	Association with positive diagnosis, prognosis or outcome

study within an established trial. Here, a two-phase process would involve an initial phase utilising more than two-thirds of the final cohort data (training cohort) to identify the most promising feature(s) and a subsequent phase using the remaining patients (independent cohort) to evaluate the performance of the identified radiomic signature. *Thirdly*, where a previously validated radiomic signature is used, this could be incorporated into a clinical trial as a primary or secondary endpoint. In this last case, the pathway of a data-driven biomarker does not differ from a QIB.

## Summary and future perspective

Data-driven imaging biomarkers provide information beyond that perceived by human readers. Their benefits may be exploited if specific standardisation and validation pathways are defined and the different/additional hurdles compared to more traditional QIBs are addressed. Effects of different types of processing on subsequent extracted feature variability and predictive model performance is an open area of research [13]. Availability of public access patient cohorts with well-documented image datasets is expected to facilitate consensus regarding pre- and post-processing methods and determine utility of radiomics within clinical trials.

While radiomics may eventually encompass all quantitative image-derived information into a common framework, current implementations mostly relate to intensity, shape and textural features within a VOI. In the future, quantitative (or even qualitative) functional information, e.g. derived from PET, SPECT, pharmacokinetic modelling and other parametric imaging modalities, may form part of the radiomic signature, and require a smaller or biologically more meaningful set of parameters. Deep radiomics may also be deployed in trials, and recent studies have already demonstrated the potential of such approaches [108–111].

Regardless of definitive biological correlation, once adopted and properly deployed, data-driven biomarkers may be combined with clinical data and other biomarkers (biochemical, genetic, epigenetic, transcription factors, proteins). Such expanded use of radiomics should eventually improve disease characterisation, prognostic stratification and response prediction in clinical trials, ultimately advancing precision medicine.

**Acknowledgements** This paper was endorsed by the ESR Executive Council in December 2020.

**Funding** The authors state that this work has not received any funding.

## Compliance with ethical standards

**Guarantor** Nandita M deSouza.

### Conflict of interest

LF - Speaker fees from Sanofi, Novartis, Janssen, General Electric.

Congress sponsorship from Guerbet. Industrial grant on radiomics from Invectys, Novartis. Co-investigator in grant with Philips, Ariana Pharma, Evolucare.

**CC**—personal fees from Pfizer, BMS, MSD, Roche and advisory role for Astra Zeneca.

**CMD** - Consulting or advisory roles with Ipsen, Novartis, Terumo, and Advanced Accelerator Applications; participation in speakers' bureaus with Terumo and Advanced Accelerator Applications; and travel, accommodations, or expenses with General Electric and Terumo.

**XG**: CEO of Gold Standard Phantoms, a company designing calibration devices for quantitative MRI.

**All other authors**- none.

**Statistics and biometry** No complex statistical methods were necessary for this paper.

**Informed consent** Not applicable in this perspectives paper.

**Ethical approval** Not applicable in this special report.

### Methodology

• Special report

**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/>.

## References

1. Santamaria G, Velasco M, Bargallo X, Caparros X, Farrus B, Luis Fernandez P (2010) Radiologic and pathologic findings in breast tumors with high signal intensity on T2-weighted MR images. *Radiographics* 30:533–548
2. Parghane RV, Basu S (2020) PET/computed tomography in treatment response assessment in cancer: an overview with emphasis on the evolving role in response evaluation to immunotherapy and radiation therapy. *PET Clin* 15:101–123
3. Lee SH, Moon WK, Cho N et al (2014) Shear-wave elastographic features of breast cancers: comparison with mechanical elasticity and histopathologic characteristics. *Invest Radiol* 49:147–155
4. de Bazelaire C, Calmon R, Chapellier M, Pluvinage A, Fria J, de Kerviler E (2010) CT and MRI imaging in tumoral angiogenesis. *Bull Cancer* 97:79–90
5. Ammari S, Thiam R, Cuenod CA et al (2014) Radiological evaluation of response to treatment: application to metastatic renal cancers receiving anti-angiogenic treatment. *Diagn Interv Imaging* 95:527–539
6. O'Connor JP, Aboagye EO, Adams JE et al (2017) Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 14:169–186
7. deSouza NM, Achten E, Alberich-Bayarri A et al (2019) Validated imaging biomarkers as decision-making tools in clinical




- trials and routine practice: current status and recommendations from the EIBALL\* subcommittee of the European Society of Radiology (ESR). *Insights Imaging* 10:87
8. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. *Radiology* 278:563–577
  9. Pinto Dos Santos D, Dietzel M, Baessler B (2020) A decade of radiomics research: are images really data or just patterns in the noise? *Eur Radiol*. <https://doi.org/10.1007/s00330-020-07108-w>
  10. Lambin P, Leijenaar RTH, Deist TM et al (2017) Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 14:749–762
  11. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B (2020) Radiomics in medical imaging—“how-to” guide and critical reflection. *Insights Imaging* 11:91
  12. Sanduleanu S, Woodruff HC, de Jong EEC et al (2018) Tracking tumor biology with radiomics: a systematic review utilizing a radiomics quality score. *Radiother Oncol* 127:349–360
  13. Nie K, Al-Hallaq H, Li XA et al (2019) NCTN assessment on current applications of radiomics in oncology. *Int J Radiat Oncol Biol Phys* 104:302–315
  14. Wang T, Gao T, Yang J et al (2019) Preoperative prediction of pelvic lymph nodes metastasis in early-stage cervical cancer using radiomics nomogram developed based on T2-weighted MRI and diffusion-weighted imaging. *Eur J Radiol* 114:128–135
  15. Cameron A, Khalvati F, Haider MA, Wong A (2016) MAPS: a quantitative radiomics approach for prostate cancer detection. *IEEE Trans Biomed Eng* 63:1145–1156
  16. Ma X, Shen F, Jia Y, Xia Y, Li Q, Lu J (2019) MRI-based radiomics of rectal cancer: preoperative assessment of the pathological features. *BMC Med Imaging* 19:86
  17. Yun J, Park JE, Lee H, Ham S, Kim N, Kim HS (2019) Radiomic features and multilayer perceptron network classifier: a robust MRI classification strategy for distinguishing glioblastoma from primary central nervous system lymphoma. *Sci Rep* 9:5746
  18. Shi L, He Y, Yuan Z et al (2018) Radiomics for response and outcome assessment for non-small cell lung cancer. *Technol Cancer Res Treat* 17:1533033818782788
  19. Peeken JC, Bernhofer M, Wiestler B et al (2018) Radiomics in radiooncology - challenging the medical physicist. *Phys Med* 48: 27–36
  20. Reuze S, Schemberg A, Orlhac F et al (2018) Radiomics in nuclear medicine applied to radiation therapy: methods, pitfalls, and challenges. *Int J Radiat Oncol Biol Phys* 102:1117–1142
  21. Elhalawani H, Lin TA, Volpe S et al (2018) Machine learning applications in head and neck radiation oncology: lessons from open-source radiomics challenges. *Front Oncol* 8:294
  22. Bibault JE, Xing L, Giraud P et al (2020) Radiomics: a primer for the radiation oncologist. *Cancer Radiother*. <https://doi.org/10.1016/j.canrad.2020.01.011>
  23. El Naqa I, Ten Haken RK (2018) Can radiomics personalise immunotherapy? *Lancet Oncol* 19:1138–1139
  24. Sun R, Limkin EJ, Vakalopoulou M et al (2018) A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* 19:1180–1191
  25. Trebeschi S, Drago SG, Birkbak NJ et al (2019) Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol* 30:998–1004
  26. Basler L, Gabrys HS, Hogan SA et al (2020) Radiomics, tumor volume and blood biomarkers for early prediction of pseudoprogression in metastatic melanoma patients treated with immune checkpoint inhibition. *Clin Cancer Res*. <https://doi.org/10.1158/1078-0432.CCR-20-0020>
  27. Choe J, Lee SM, Do KH et al (2020) Outcome prediction in resectable lung adenocarcinoma patients: value of CT radiomics. *Eur Radiol*. <https://doi.org/10.1007/s00330-020-06872-z>
  28. Capobianco E, Dominietto M (2020) From medical imaging to radiomics: role of data science for advancing precision health. *J Pers Med* 10
  29. Bogowicz M, Riesterer O, Ikenberg K et al (2017) Computed tomography radiomics predicts HPV status and local tumor control after definitive radiochemotherapy in head and neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 99:921–928
  30. Zhong Y, Yuan M, Zhang T, Zhang YD, Li H, Yu TF (2018) Radiomics approach to prediction of occult mediastinal lymph node metastasis of lung adenocarcinoma. *AJR Am J Roentgenol* 211:109–113
  31. Dou TH, Coroller TP, van Griethuysen JJM, Mak RH, Aerts H (2018) Peritumoral radiomics features predict distant metastasis in locally advanced NSCLC. *PLoS One* 13:e0206108
  32. Kim JY, Park JE, Jo Y et al (2019) Incorporating diffusion- and perfusion-weighted MRI into a radiomics model improves diagnostic performance for pseudoprogression in glioblastoma patients. *Neuro Oncol* 21:404–414
  33. Suh HB, Choi YS, Bae S et al (2018) Primary central nervous system lymphoma and atypical glioblastoma: differentiation using radiomics approach. *Eur Radiol* 28:3832–3839
  34. Li Y, Liu X, Xu K et al (2018) MRI features can predict EGFR expression in lower grade gliomas: a voxel-based radiomic analysis. *Eur Radiol* 28:356–362
  35. Li Y, Qian Z, Xu K et al (2018) MRI features predict p53 status in lower-grade gliomas via a machine-learning approach. *Neuroimage Clin* 17:306–311
  36. Wormald BW, Doran SJ, Ind TE, D’Arcy J, Petts J, deSouza NM (2020) Radiomic features of cervical cancer on T2- and diffusion-weighted MRI: prognostic value in low-volume tumors suitable for trachelectomy. *Gynecol Oncol* 156:107–114
  37. Cook GJ, Yip C, Siddique M et al (2013) Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med* 54:19–26
  38. Tixier F, Hatt M, Valla C et al (2014) Visual versus quantitative assessment of intratumor 18F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. *J Nucl Med* 55: 1235–1241
  39. Cook GJ, O’Brien ME, Siddique M et al (2015) Non-small cell lung cancer treated with erlotinib: heterogeneity of (18)F-FDG uptake at PET-association with treatment response and prognosis. *Radiology* 276:883–893
  40. Parmar C, Leijenaar RT, Grossmann P et al (2015) Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep* 5:11044
  41. Li S, Ding C, Zhang H, Song J, Wu L (2019) Radiomics for the prediction of EGFR mutation subtypes in non-small cell lung cancer. *Med Phys* 46:4545–4552
  42. Mattonen SA, Davidzon GA, Benson J et al (2019) Bone marrow and tumor radiomics at (18)F-FDG PET/CT: impact on outcome prediction in non-small cell lung cancer. *Radiology* 293:451–459
  43. Antunes J, Viswanath S, Rusu M et al (2016) Radiomics analysis on FLT-PET/MRI for characterization of early treatment response in renal cell carcinoma: a proof-of-concept study. *Transl Oncol* 9: 155–162
  44. Zamboglou C, Carles M, Fechter T et al (2019) Radiomic features from PSMA PET for non-invasive intraprostatic tumor discrimination and characterization in patients with intermediate- and high-risk prostate cancer - a comparison study with histology reference. *Theranostics* 9:2595–2605
  45. Zheng X, Yao Z, Huang Y et al (2020) Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun* 11:1236

46. Caramella C, Allorant A, Orlhac F et al (2018) Can we trust the calculation of texture indices of CT images? A phantom study. *Med Phys* 45:1529–1536
47. Raunig DL, McShane LM, Pennello G et al (2015) Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. *Stat Methods Med Res* 24:27–67
48. Shaikh F, Franc B, Allen E et al (2018) Translational radiomics: defining the strategy pipeline and considerations for application-part 2: from clinical implementation to enterprise. *J Am Coll Radiol* 15:543–549
49. Shaikh F, Franc B, Allen E et al (2018) Translational radiomics: defining the strategy pipeline and considerations for application-part 1: from methodology to clinical implementation. *J Am Coll Radiol* 15:538–542
50. Zhao B, Tan Y, Tsai WY et al (2016) Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep* 6:23428
51. Pfähler E, van Sluis J, Merema BBJ et al (2020) Experimental multicenter and multivendor evaluation of the performance of PET radiomic features using 3-dimensionally printed phantom inserts. *J Nucl Med* 61:469–476
52. Mackin D, Fave X, Zhang L et al (2017) Harmonizing the pixel size in retrospective computed tomography radiomics studies. *PLoS One* 12:e0178524
53. Nyul LG, Udupa JK, Zhang X (2000) New variants of a method of MRI scale standardization. *IEEE Trans Med Imaging* 19:143–150
54. Isaksson LJ, Raimondi S, Botta F et al (2020) Effects of MRI image normalization techniques in prostate cancer radiomics. *Phys Med* 71:7–13
55. Scalco E, Belfatto A, Mastropietro A et al (2020) T2w-MRI signal normalization affects radiomics features reproducibility. *Med Phys* 47:1680–1691
56. Leijenaar RT, Nalbantov G, Carvalho S et al (2015) The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep* 5:11075
57. Lee SH, Cho HH, Lee HY, Park H (2019) Clinical impact of variability on CT radiomics and suggestions for suitable feature selection: a focus on lung cancer. *Cancer Imaging* 19:54
58. Duron L, Balvay D, Vande Perre S et al (2019) Gray-level discretization impacts reproducible MRI radiomics texture features. *PLoS One* 14:e0213459
59. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I (2019) Validation of a method to compensate multicenter effects affecting CT radiomics. *Radiology* 291:53–59
60. Rogers W, Thulasi Seetha S, Refaee TAG et al (2020) Radiomics: from qualitative to quantitative imaging. *Br J Radiol* 93:20190948
61. Zwanenburg A, Vallières M, Abdalah MA et al (2020) The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* 295:328–338
62. Owens CA, Peterson CB, Tang C et al (2018) Lung tumor segmentation methods: impact on the uncertainty of radiomics features for non-small cell lung cancer. *PLoS One* 13:e0205003
63. Caballo M, Pangallo DR, Mann RM, Sechopoulos I (2020) Deep learning-based segmentation of breast masses in dedicated breast CT imaging: radiomic feature stability between radiologists and artificial intelligence. *Comput Biol Med* 118:103629
64. Hatt M, Lee JA, Schmidlein CR et al (2017) Classification and evaluation strategies of auto-segmentation approaches for PET: report of AAPM task group No. 211. *Med Phys* 44:e1–e42
65. Waninger JJ, Green MD, Cheze Le Rest C, Rosen B, El Naqa I (2019) Integrating radiomics into clinical trial design. *Q J Nucl Med Mol Imaging* 63:339–346
66. Ciardo D, Jereczek-Fossa BA, Petralia G et al (2017) Multimodal image registration for the identification of dominant intraprostatic lesion in high-precision radiotherapy treatments. *Br J Radiol* 90:20170021
67. Ou Y, Weinstein SP, Conant EF et al (2015) Deformable registration for quantifying longitudinal tumor changes during neoadjuvant chemotherapy. *Magn Reson Med* 73:2343–2356
68. Fornaçon-Wood I, Mistry H, Ackermann CJ et al (2020) Reliability and prognostic value of radiomic features are highly dependent on choice of feature extraction platform. *Eur Radiol* 30:6241–6250
69. Dhali DKR, Juneja M (2020) Machine learning: a review of the algorithms and its applications. In: Singh PKASY, Kolekar M, Tanwar S (eds) *Proceedings of ICRIC 2019 Lecture Notes in Electrical Engineering*. Springer, Cham
70. Ozgur C, Kleckner M, Li Y (2015) Selection of statistical software for solving big data problems: a guide for businesses, students, and universities. *Sage Open* 5:1–12
71. Pillai R, Oza P, Sharma P (2020) Review of machine learning techniques in health care. In: Singh P, Kar A, Singh Y, Kolekar M, Tanwar S (eds) *Proceedings of ICRIC 2019 Lecture Notes in Electrical Engineering*. Springer, Cham
72. Tanwani AK, Alfridi J, Shafiq Z, Farooq M (2009) Guidelines to select machine learning scheme for classification of biomedical datasets. In: Pizzuti C, Ritchie MD, Giacobini M (eds) *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics EvoBIO 2009 Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, pp 128–139
73. Chen T, Ning Z, Xu L et al (2019) Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumours preoperatively. *Eur Radiol* 29:1074–1082
74. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts H (2015) Machine Learning methods for quantitative radiomic biomarkers. *Sci Rep* 5:13087
75. Leger S, Zwanenburg A, Pilz K et al (2017) A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep* 7:13206
76. Afshar P, Mohammadi A, Plataniotis KN, Oikonomou A, Benali H (2019) From handcrafted to deep-learning-based cancer radiomics challenges and opportunities. *Ieee Signal Processing Magazine* 36:132–160
77. Vial A, Stirling D, Field M et al (2018) The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review. *Transl Cancer Res* 7:803–816
78. Avanzo M, Wei L, Stancanello J et al (2020) Machine and deep learning methods for radiomics. *Med Phys* 47:e185–e202
79. Peerlings J, Woodruff HC, Winfield JM et al (2019) Stability of radiomics features in apparent diffusion coefficient maps from a multi-centre test-retest trial. *Sci Rep* 9:4800
80. Traverso A, Wee L, Dekker A, Gillies R (2018) Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol Biol Phys* 102:1143–1158
81. AlBadawy EA, Saha A, Mazurowski MA (2018) Deep learning for segmentation of brain tumors: impact of cross-institutional training and testing. *Med Phys* 45:1150–1158
82. Larue R, Klaassen R, Jochems A et al (2018) Pre-treatment CT radiomics to predict 3-year overall survival following chemoradiotherapy of esophageal cancer. *Acta Oncol* 57:1475–1481
83. Soufi M, Arimura H, Nagami N (2018) Identification of optimal mother wavelets in survival prediction of lung cancer patients using wavelet decomposition-based radiomic features. *Med Phys* 45:5116–5128
84. Xu X, Huang L, Chen J et al (2019) Application of radiomics signature captured from pretreatment thoracic CT to predict brain metastases in stage III/IV ALK-positive non-small cell lung cancer patients. *J Thorac Dis* 11:4516–4528
85. Li H, Xie Y, Wang X, Chen F, Sun J, Jiang X (2019) Radiomics features on non-contrast computed tomography predict early

- enlargement of spontaneous intracerebral hemorrhage. *Clin Neurol Neurosurg* 185:105491
86. Huynh E, Coroller TP, Narayan V et al (2016) CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. *Radiother Oncol* 120:258–266
  87. Hui B, Qiu JJ, Liu JH, Ke NW (2020) Identification of pancreaticoduodenectomy resection for pancreatic head adenocarcinoma: a preliminary study of radiomics. *Comput Math Methods Med* 2020:2761627
  88. Leithner D, Horvat JV, Marino MA et al (2019) Radiomic signatures with contrast-enhanced magnetic resonance imaging for the assessment of breast cancer receptor status and molecular subtypes: initial results. *Breast Cancer Res* 21:106
  89. Zhang Y, Yan P, Liang F, Ma C, Liang S, Jiang C (2019) Predictors of epilepsy presentation in unruptured brain arteriovenous malformations: a quantitative evaluation of location and radiomics features on T2-weighted imaging. *World Neurosurg* 125:e1008–e1015
  90. Zhou J, Lu J, Gao C et al (2020) Predicting the response to neo-adjuvant chemotherapy for breast cancer: wavelet transforming radiomics in MRI. *BMC Cancer* 20:100
  91. Lue KH, Wu YF, Liu SH et al (2019) Intratumor heterogeneity assessed by (18)F-FDG PET/CT predicts treatment response and survival outcomes in patients with Hodgkin lymphoma. *Acad Radiol*. <https://doi.org/10.1016/j.acra.2019.10.015>
  92. Shiri I, Maleki H, Hajianfar G et al (2020) Next-generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using multimodal imaging and machine learning algorithms. *Mol Imaging Biol*. <https://doi.org/10.1007/s11307-020-01487-8>
  93. Lee SH, Han P, Hales R et al (2020) Multi-view radiomics and dosiomics analysis with machine learning for predicting acute-phase weight loss in lung cancer patients treated with radiotherapy. *Phys Med Biol*. <https://doi.org/10.1088/1361-6560/ab8531>
  94. Nazari M, Shiri I, Hajianfar G et al (2020) Noninvasive Fuhman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning. *Radiol Med*. <https://doi.org/10.1007/s11547-020-01169-z>
  95. Bhatia A, Birger M, Veeraraghavan H et al (2019) MRI radiomic features are associated with survival in melanoma brain metastases treated with immune checkpoint inhibitors. *Neuro Oncol* 21:1578–1586
  96. Shayesteh SP, Alikhassi A, Fard Esfahani A et al (2019) Neo-adjuvant chemoradiotherapy response prediction using MRI based ensemble learning method in rectal cancer patients. *Phys Med* 62:111–119
  97. Fiset S, Welch ML, Weiss J et al (2019) Repeatability and reproducibility of MRI-based radiomic features in cervical cancer. *Radiother Oncol* 135:107–114
  98. Fave X, Zhang L, Yang J et al (2017) Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Sci Rep* 7:588
  99. Dong X, Sun X, Sun L et al (2016) Early change in metabolic tumor heterogeneity during chemoradiotherapy and its prognostic value for patients with locally advanced non-small cell lung cancer. *PLoS One* 11:e0157836
  100. Tixier F, Vriens D, Cheze-Le Rest C et al (2016) Comparison of tumor uptake heterogeneity characterization between static and parametric 18F-FDG PET images in non-small cell lung cancer. *J Nucl Med* 57:1033–1039
  101. Yin P, Mao N, Zhao C, Wu J, Chen L, Hong N (2019) A triple-classification radiomics model for the differentiation of primary chordoma, giant cell tumor, and metastatic tumor of sacrum based on T2-weighted and contrast-enhanced T1-weighted MRI. *J Magn Reson Imaging* 49:752–759
  102. Grossmann P, Stringfield O, El-Hachem N et al (2017) Defining the biological basis of radiomic phenotypes in lung cancer. *Elife* 6
  103. Tu W, Sun G, Fan L et al (2019) Radiomics signature: a potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. *Lung Cancer* 132:28–35
  104. Sanghani P, Ang BT, King NKK, Ren H (2018) Overall survival prediction in glioblastoma multiforme patients from volumetric, shape and texture features using machine learning. *Surg Oncol* 27:709–714
  105. Ferreira Junior JR, Koenigkam-Santos M, Cipriano FEG, Fabro AT, Azevedo-Marques PM (2018) Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. *Comput Methods Programs Biomed* 159:23–30
  106. Penzias G, Singanamalli A, Elliott R et al (2018) Identifying the morphologic basis for radiomic features in distinguishing different Gleason grades of prostate cancer on MRI: preliminary findings. *PLoS One* 13:e0200730
  107. Holm EA (2019) In defense of the black box. *Science* 364:26–27
  108. Oakden-Rayner L, Carneiro G, Bessen T, Nascimento JC, Bradley AP, Palmer LJ (2017) Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Sci Rep* 7:1648
  109. Bibault JE, Giraud P, Housset M et al (2018) Deep learning and radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. *Sci Rep* 8:12611
  110. Ning Z, Luo J, Li Y et al (2019) Pattern classification for gastrointestinal stromal tumors by integration of radiomics and deep convolutional features. *IEEE J Biomed Health Inform* 23:1181–1191
  111. Shboul ZA, Alam M, Vidyaratne L, Pei L, Elbakary MI, Iftekharuddin KM (2019) Feature-guided deep radiomics for glioblastoma patient survival prediction. *Front Neurosci* 13:966

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Laure Fournier<sup>1,2,3</sup> · Lena Costaridou<sup>2,4</sup> · Luc Bidaut<sup>3,5</sup> · Nicolas Michoux<sup>3,6</sup> · Frederic E. Lecouvet<sup>3,6</sup> · Lioe-Fee de Geus-Oei<sup>3,7,8</sup> · Ronald Boellaard<sup>2,9,10</sup> · Daniela E. Oprea-Lager<sup>3,9</sup> · Nancy A Obuchowski<sup>10,11</sup> · Anna Caroli<sup>2,12</sup> · Wolfgang G. Kunz<sup>3,13</sup> · Edwin H. Oei<sup>2,14</sup> · James P. B. O'Connor<sup>2,15</sup> · Marius E. Mayerhoefer<sup>2,16</sup> · Manuela Franca<sup>2,17</sup> · Angel Alberich-Bayarri<sup>2,18</sup> · Christophe M. Deroose<sup>3,19,20</sup> · Christian Loewe<sup>2,21</sup> · Rashindra Manniesing<sup>2,22</sup> · Caroline Caramella<sup>3,23</sup> · Egesta Lopci<sup>3,24</sup> · Nathalie Lassau<sup>2,3,10,25</sup> · Anders Persson<sup>2,26</sup> · Rik Achten<sup>2,27</sup> · Karen Rosendahl<sup>2,28</sup> · Olivier Clement<sup>1,2</sup> · Elmar Kotter<sup>2,29</sup> · Xavier Golay<sup>2,10,30</sup> · Marion Smits<sup>2,3,14</sup> · Marc Dewey<sup>2,31</sup> · Daniel C. Sullivan<sup>2,10,32</sup> · Aad van der Lugt<sup>2,14</sup> · Nandita M. deSouza<sup>2,3,10,33</sup>  · European Society of Radiology<sup>34</sup>

<sup>1</sup> PARCC, INSERM, Radiology Department, AP-HP, Hopital europeen Georges Pompidou, Université de Paris, F-75015 Paris, France

<sup>2</sup> European Imaging Biomarkers Alliance (EIBALL), European Society of Radiology, Vienna, Austria

<sup>3</sup> Imaging Group, European Organisation of Research and Treatment in Cancer (EORTC), Brussels, Belgium

<sup>4</sup> School of Medicine, University of Patras, University Campus, Rio, 26 500 Patras, Greece

<sup>5</sup> College of Science, University of Lincoln, Lincoln LN6 7TS, UK

<sup>6</sup> Department of Radiology, Institut de Recherche Expérimentale et Clinique (IREC), Cliniques Universitaires Saint Luc, Université Catholique de Louvain (UCLouvain), B-1200 Brussels, Belgium

<sup>7</sup> Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands

<sup>8</sup> Biomedical Photonic Imaging Group, University of Twente, Enschede, The Netherlands

<sup>9</sup> Department of Radiology & Nuclear Medicine, Cancer Centre Amsterdam, Amsterdam University Medical Centers (VU University), Amsterdam, The Netherlands

<sup>10</sup> Quantitative Imaging Biomarkers Alliance, Radiological Society of North America, Oak Brook, IL, USA

<sup>11</sup> Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, USA

<sup>12</sup> Department of Biomedical Engineering, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Bergamo, Italy

<sup>13</sup> Department of Radiology, University Hospital, LMU Munich, Munich, Germany

<sup>14</sup> Department of Radiology & Nuclear Medicine, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

<sup>15</sup> Division of Cancer Sciences, University of Manchester, Manchester, UK

<sup>16</sup> Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria

<sup>17</sup> Department of Radiology, Centro Hospitalar Universitário do Porto, Instituto de Ciências Biomédicas de Abel Salazar, University of Porto, Porto, Portugal

<sup>18</sup> Quantitative Imaging Biomarkers in Medicine (QUIBIM), Valencia, Spain

<sup>19</sup> Nuclear Medicine, University Hospitals Leuven, Leuven, Belgium

<sup>20</sup> Nuclear Medicine and Molecular Imaging, Department of Imaging and Pathology, KU Leuven, Leuven, Belgium

<sup>21</sup> Division of Cardiovascular and Interventional Radiology, Dept. for Bioimaging and Image-Guided Therapy, Medical University of Vienna, Vienna, Austria

<sup>22</sup> Department of Radiology and Nuclear Medicine, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands

<sup>23</sup> Radiology Department, Hôpital Marie Lannelongue, Institut d'Oncologie Thoracique, Université Paris-Saclay, Le Plessis-Robinson, France

<sup>24</sup> Nuclear Medicine, Humanitas Clinical and Research Hospital – IRCCS, Rozzano, MI, Italy

<sup>25</sup> Imaging Department, Gustave Roussy Cancer Campus Grand, Paris, UMR 1281, INSERM, CNRS, CEA, Université Paris-Saclay, Saint-Aubin, France

<sup>26</sup> Department of Radiology, and Department of Health, Medicine and Caring Sciences, Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden

<sup>27</sup> Department of Radiology and Medical Imaging, Ghent University Hospital, Ghent, Belgium

<sup>28</sup> Department of Radiology, University Hospital of North Norway, Tromsø, Norway

<sup>29</sup> Department of Radiology, University Medical Center Freiburg, Freiburg, Germany

<sup>30</sup> Queen Square Institute of Neurology, University College London, London, UK

<sup>31</sup> Department of Radiology, Charité Universitätsmedizin Berlin, Berlin, Germany

<sup>32</sup> Dept. of Radiology, Duke University, 311 Research Dr, Durham, NC 27710, USA

<sup>33</sup> Division of Radiotherapy and Imaging, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK

<sup>34</sup> Am Gestade 1, 1010 Vienna, Austria