Radiomics: the bridge between medical imaging and personalized medicine

Philippe Lambin[1], Ralph T.H. Leijenaar[1], Timo M. Deist[1]*, Jurgen Peerlings[1,2], Evelyn E.C. de Jong[1], Janita van Timmeren[1], Sebastian Sanduleanu[1], Ruben T.H.M. Larue[1], Aniek J.G. Even[1], Arthur Jochems[1], Yvonka van Wijk[1], Henry Woodruff[1], Johan van Soest[3], Tim Lustberg[3], Erik Roelofs[1,3], Wouter van Elmpt[3], Andre Dekker[3], Felix M. Mottaghy[2,4], Joachim E. Wildberger[2] and Sean Walsh[1]*

Abstract | Radiomics, the high-throughput mining of quantitative image features from standard-of-care medical imaging that enables data to be extracted and applied within clinical-decision support systems to improve diagnostic, prognostic, and predictive accuracy, is gaining importance in cancer research. Radiomic analysis exploits sophisticated image analysis tools and the rapid development and validation of medical imaging data that uses image-based signatures for precision diagnosis and treatment, providing a powerful tool in modern medicine. Herein, we describe the process of radiomics, its pitfalls, challenges, opportunities, and its capacity to improve clinical decision making, emphasizing the utility for patients with cancer. Currently, the field of radiomics lacks standardized evaluation of both the scientific integrity and the clinical relevance of the numerous published radiomics investigations resulting from the rapid growth of this area. Rigorous evaluation criteria and reporting guidelines need to be established in order for radiomics to mature as a discipline. Herein, we provide guidance for investigations to meet this urgent need in the field of radiomics.

[1] *The D-Lab: Decision Support for Precision Medicine, GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, Universiteitssingel 40, 6229 ER, Maastricht, The Netherlands.*
[2] *Department of Radiology and Nuclear Medicine, GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, Doctor Tanslaan 12, 6229 ET, Maastricht, The Netherlands.*
[3] *Department of Radiation Oncology (MAASTRO), GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, Doctor Tanslaan 12, 6229 ET, Maastricht, The Netherlands.*
[4] *Department of Nuclear Medicine, University Hospital RWTH Aachen, Pauwelsstraße 30, 52074 Aachen, Germany.*
**These authors contributed equally to this work.*
Correspondence to P. L. philippe.lambin@maastrichtuniversity.nl

doi:10.1038/nrclinonc.2017.141
Published online 4 Oct 2017

Imaging is an important technology in medical science and is used in clinical practice to aid decision making[1]. The role of medical imaging, however, is swiftly evolving from being primarily a diagnostic tool to also include a central role in the context of personalized precision medicine[2]. In radiomics[3,4], digitally encrypted medical images that hold information related to tumour pathophysiology are transformed into mineable high-dimensional data[1]. This information can be harnessed through quantitative image analyses[5] and leveraged via clinical-decision support systems (CDSS)[6] to improve medical decision-making. Radiomics builds upon several decades of computer-aided diagnosis, prognosis, and therapeutics research[7,8]. The process used in radiomics involves the identification of vast arrays of quantitative features within digital images, storage of such data in federated databases (that is, a system in which several independent databases function as a single entity) and the subsequent mining of the data for knowledge extraction and application[9]. Innumerable quantitative features can now be extracted using high-throughput computing from medical images such as CT, MR, and/or PET. The creation of databases that link immense volumes of radiomics data (ideally with all other pertinent data) from millions of patients to form vast, rapid learning healthcare (RLHC) networks is conceivable, but presents a considerable data management hurdle[10–13].

Radiomics is not a panacea for clinical decision-making. Radiomic features (such as intensity, shape, texture or wavelet) offer information on cancer phenotype as well as the tumour microenvironment that is distinct and complementary to other pertinent data sources (including clinically obtained, treatment-related or genomic data)[14]. Radiomics-derived data, when combined with other pertinent data and correlated and/or inferred with outcomes data, can produce accurate robust evidence-based CDSS.

The potential of radiomics to improve CDSS is beyond doubt[15] and the field is evolving rapidly. The principal challenge is the optimal collection and integration of diverse multimodal data sources in a quantitative manner that delivers unambiguous clinical predictions that accurately and robustly enable outcome prediction as a function of the impending decisions[16]. Many published prediction models that account for factors related to both disease and treatment are available, but these models lack standardized

Key points

- Radiomics is becoming increasingly more important in medical imaging
- The explosion of medical imaging data creates an environment ideal for machine-learning and data-based science
- Radiomics-based decision-support systems for precision diagnosis and treatment can be a powerful tool in modern medicine
- Large-scale data sharing is necessary for the validation and full potential that radiomics represents
- Standardized data collection, evaluation criteria, and reporting guidelines are required for radiomics to mature as a discipline

evaluation of their performance, reproducibility, and/or clinical utility[17]. Consequently, these models might not be appropriate for CDSS.

In this Review, we describe the process of radiomics along with latest developments in the field. The pitfalls, challenges, and opportunities presented by radiomics to improve CDSS for personalized precision oncology are highlighted, with an emphasis on the methodological aspects of radiomics prediction model development and validation. We explore the advanced and innovative information technologies that are essential for the data management of diverse multimodal data sources. Finally, we offer a vision of the necessary steps to ensure continued progression and widespread acceptance of both radiomics and CDSS.

The workflow of radiomics

Radiomics is defined as the quantitative mapping, that is, extraction, analysis and modelling of many medical image features in relation to prediction targets, such as clinical end points and genomic features. A radiomics study can be structured in five phases: data selection, medical imaging, feature extraction, exploratory analysis, and modelling (FIG. 1). To assess the quality of radiomics studies, we propose the radiomics quality score (RQS).

Data selection

Radiomic analyses begins with the choice of an imaging protocol, the volume of interest (VOI) and a prediction target — the event one wishes to predict. Typically, the entire primary tumour is analysed and linked to available data on treatment outcomes, such as survival. Radiomic analyses can be performed on subregions of the tumour (habitats), metastatic lesions, as well as in normal tissues. Analysis of these regions might yield radiosensitive phenotypes, which has implications for treatment planning strategies. Radiomics analysis, however, is not restricted to radiotherapy and can be applied to any image generated in the clinical setting (FIG. 2).

The importance of using standardized imaging protocols to eliminate unnecessary confounding variability is recognized[9,18]; however, nonstandardized imaging protocols are commonplace. Therefore, reproducibility and

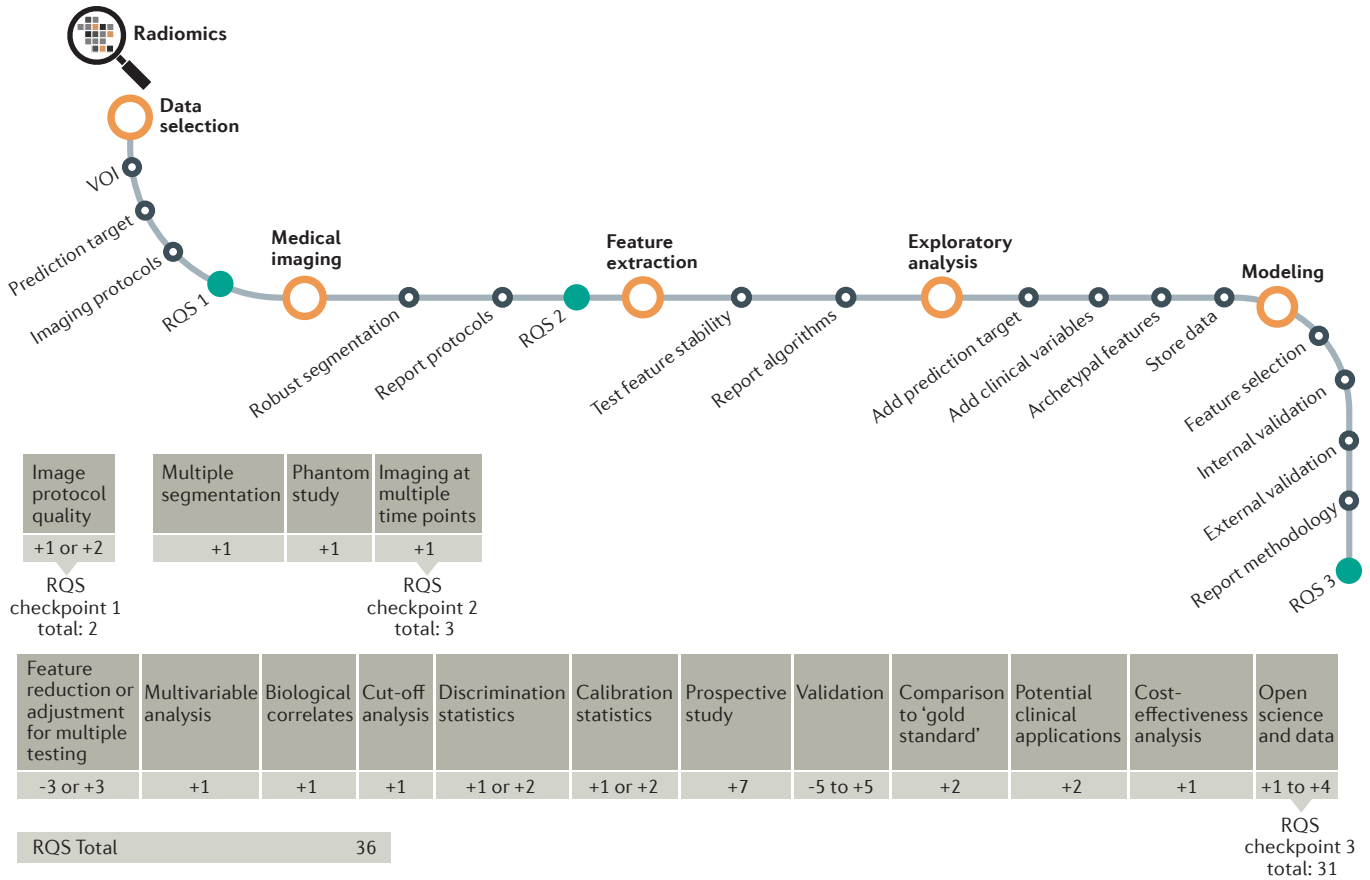

Figure 1 | **Flowchart depicting the workflow of radiomics and the application of the RQS.** The workflow includes the necessary steps in a radiomic analysis. The RQS both rewards and penalizes the methodology and analyses of a study, consequently encouraging the best scientific practice. RSQ, radiomics quality score; VOI, volume of interest.

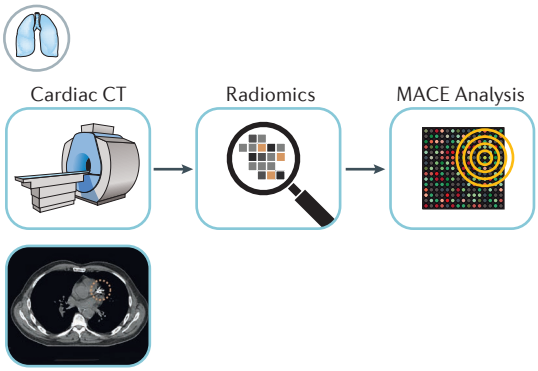

Figure 2 | **Radiomics in cardiology.** The current gold standard for quantification of coronary calcifications visible on CT is the 'Agatston' method (based upon intensity and volume). Radiomic features can improve quantification, differentiation between calcified and non-calcified plaques, and thus the prediction of Major Adverse Cardiac Events (MACE).

comparability of radiomic studies can be achieved only by extensive disclosure of imaging protocols. We wish to emphasize this point, and provide examples of how protocols should be reported in future radiomics studies (Supplementary information S1).

Medical imaging

Segmentation. VOIs are segmented manually or (semi-) automatically[19]. This segmentation determines which voxels within an image are analysed, thus, the variability in segmentation can introduce bias in the evaluation of derived radiomic features[20]. Multiple-segmentation is a method to limit the extent of this bias. Examples that enable robust features to be observed[21] include: evaluation by multiple clinicians, perturb segmentations with noise, combination of diverse algorithms, or use different stages of the breathing cycle. Key considerations are how the segmentation was performed, and how sensitive the radiomics analysis is to different segmentation methods[22]. For example, a semi-automatic segmentation method can result in different radiomic features than a manual delineation.

Phantom studies. The determination of inter-scanner and inter-vendor variability of features is important in radiomics[23]. In cases in which radiomic studies rely on data from multiple scanners, neglecting this variability can jeopardize the analysis of studies — that is, the proposed radiomic-based prediction model might not perform adequately on external datasets if new data are acquired on different scanners. As data from patients scanned on multiple devices is scarce and subject to uncertainties (such as organ motion, or different imaging protocols), phantom studies are a suitable means to gauge these uncertainties and identify features that rely on the vendor. In essence, phantom studies provide a risk-mitigation strategy to help navigate from the current clinical imaging scenario to the desired optimal imaging scenario.

Imaging at multiple time points. Additional sources of variability in radiomics features are organ motion or expansion or shrinkage of the target volume. Radiomics features that are strongly dependent on these factors can have limited applicability. To account for these sources of variability, available test-retest data[24–26] can be exploited to measure radiomics feature stability. For example, two datasets of images acquired within a small period of time from a patient cohort.

Phantom studies
An artificial structure that imitates human tissue properties is scanned on multiple machines to characterize scan output against a known physical standard.

Feature extraction

The essence of radiomics is the high-throughput extraction of quantitative image features to characterize VOIs. Feature values are dependent upon factors that can include image pre-processing (for example, filtering, or intensity discretization) and reconstruction (for example, filtered back projection, or iterative reconstruction). Furthermore, variation exists in feature nomenclature, mathematical definition, methodology, and software implementation of the applied feature extraction algorithms[27–29]. In order to facilitate inter-operability of radiomic features, differences in nomenclature, algorithms, software implementations, as well as other methodological aspects must be elucidated.

Exploratory analysis

Radiomic and non-radiomic features should be combined with the prediction target to create a single dataset. This approach enables the investigation of relationships between features. Groups of highly correlated radiomics features can be identified via clustering, and these features can be reduced to single archetypal features per cluster. Radiomic features that are well-correlated with routine clinical features (such as tumour stage) do not provide additional information. Auxiliary feature data collected from multiple segmentations, multiple imaging, and phantom studies, can be exploited to assess feature robustness. Volatile or robust features can be identified and subsequently excluded from model development. For example, a feature that is robust for the prediction of overall survival for lung cancer (that is, imaged and segmented in a certain way) for a given dataset could be volatile for the prediction of pneumonitis in lung cancer (imaged and segmented in an alternative way) for a given dataset. Thus, the process of feature reduction and/or exclusion should be described clearly.

Modelling

Radiomic modelling involves three major aspects: feature selection, modelling methodology, and validation. Feature selection should be data-driven owing to the vast in-human range of possible radiomics features; such analysis should be performed in a robust and transparent manner. To achieve holistic models, features beyond radiomics (such as data from clinical records, data obtained during treatment or biological and/or genetic) should also be incorporated. Regarding the choice of modelling methodology, the identification of optimal machine-learning methods for radiomic applications is a crucial step towards stable and clinically relevant CDSS; thus, in the ideal scenario, multiple machine-learning methods should be employed[30] and the implementation should be comprehensively documented. A non-validated model is

of limited value; validation is an indispensable component of a complete radiomic analysis. Models must be internally validated and, ideally, should be externally validated.

Feature selection. Depending on the number of filters, feature categories, and other adjustable parameters, the possible number of radiomic features that can be extracted from images is virtually unlimited. The inclusion of all possible features in a model would inevitably result in overfitting, which jeopardizes model performance in patients not previously evaluated. To avoid overfitting, features that lack robustness against sources of variability should be eliminated, and archetypal features selected via dimensionality reduction techniques (such as principal component analysis or clustering). For example, a feature that is archetypal for the prediction of overall survival in patients with lung cancer for a given dataset (imaged and segmented in a certain way) could be redundant for the prediction of pneumonitis in lung cancer for a given dataset (imaged and segmented in an alternative way).

Modelling methodology. The modelling methodology chosen is often a single technique, selected according to the preference and experience of those conducting the study. Different techniques are associated with distinct inherent limitations, which include the independence assumption for features in logistic regression, the need for feature discretization in Bayesian networks, or the network configuration dependency in deep learning. The choice of modelling technique has been shown to affect prediction performance in radiomics[30]. Thus, multiple-modelling methodology implementations are desirable, but not essential. The key aspect in the selection of a modelling methodology is that, when reported, the work must be entirely reproducible. This goal can be achieved, ideally, by making the software code available (for example, via github[31]). (See Supplementary information S1 material for an overview of machine learning techniques).

Validation. Validation techniques are useful tools to assess model performance, and thus, internal and/or preferably external validation must be performed. Researchers must assess whether the model is predictive for the target patient population or just for a particular subset of samples analysed. Model performance is typically measured in terms of discrimination and calibration. Discrimination can be reported in terms of the receiver operating characteristic (ROC) curve, or the area under the ROC curve (AUC). The AUC quantifies the sensitivity and specificity of the model and represents the probability that a randomly selected patient matching an outcome is assigned that outcome by the prediction model with a larger event-probability than a randomly chosen patient who does not match the outcome. Calibration refers to the agreement between observed outcomes and model predictions, typically based on grouping of predictions. For example, the predictions are grouped according to high, medium or low probability. If the mean prediction of tumour recurrence in the high-probability group is 25%, the observed frequency of tumour recurrence in this group should ideally be 25 out of 100 patients. Calibration can be reported using a calibration plot and calibration-in-the-large/slope. A measure of overall performance is the Brier score, the mean squared prediction error. All statistical methods should be reported for training data and validation data. Valid models should exhibit statistical consistency between the training and validation sets. Bootstrapping techniques can be used to estimate confidence intervals for the abovementioned statistics and should be reported. An externally validated model has more credibility than an internally validated model, because data obtained with the former approach are considered more independent, which reinforces the validation. A large body of literature on validation techniques is available[32–35].

Reporting open-access scientific data

Validation is the first step towards a model being accepted in both the scientific and clinical communities. Independent verification of the results is a necessary additional step. Reproduction means verification of the results by independent researchers repeating the analysis using an identical technique and the same dataset and/or patient cohort, ensuring that the analysis is conducted without error. Replication means independent verification of the results by independent researchers repeating the analysis using the same technique and different (but appropriately selected) datasets and/or patient cohorts, aiming for a stronger affirmation of the findings[36–39]. Radiomic studies involve multiple complex subprocesses (such as data selection, image acquisition, feature extraction, or modelling), each one affected by a wide range of decisions, use of nonstandardized terminology, establishment of parameters, and software selection. Reproducibility and replicability in radiomics are impossible if researchers do not disclose these intricacies. The amount of necessary information far exceeds the limits of a traditional manuscript. We propose that future publications including radiomic results should provide the following as supplementary material: disclosure of imaging protocols, analysed scans, segmentations of VOIs, detailed accounts of how features were extracted (including the formulae), and of the modelling methodology used (ideally, the code). This level of meticulous detail is required in order to facilitate reproduction and replication. Furthermore, multiple radiomics software packages are available and are subject to updates or version-control. We recognize that the publication of data derived from patients might not be feasible in all circumstances. As a minimal means of comparison, and to alleviate this lack of transparency, we propose that researchers publish numerical values of their investigated features computed on the digital phantom described in the supplementary material of this manuscript (available online[40]).

To compare different software implementations for radiomic feature-extraction algorithms, we present an example, in which CT-obtained data of the primary tumour region and the corresponding tumour contours of four patients with lung cancer serve as 'real-life' digital phantoms (FIG. 3). Using the preprocessed image data, we calculated a set of commonly used features to serve as a

Calibration-in-the-large
Describes whether the predictions deviate systematically (intercept), whereas the calibration slope should ideally be equal to 1.

The independence assumption
The definition in terms of conditional probabilities is that the probability of B is not changed by knowing that A has occurred. Statistically independent variables are always uncorrelated, but the converse is not necessarily true.

Feature discretization
The process of converting continuous features to discrete binned interval features.

Bootstrapping
Measures the accuracy (defined in terms of bias, variance, confidence intervals, prediction error, etc.) to characterize the sample distribution by way of repeated random sampling methods.

reference feature dataset (See Supplementary information S1 for a detailed description of the digital phantom image data and calculated features).

The radiomics quality score

There is an urgent need for homogeneous evaluation criteria and reporting guidelines in order for radiomics to develop as a field. We propose the radiomics quality score (RQS)[41] to aid assessment of both past and future radiomic studies.

Editors, reviewers, and readers should be able to easily ascertain whether a radiomic study is compliant with best-practice procedures or, alternatively, whether the study investigators have sufficiently justified any non-compliance with guidelines. Publications should clearly state how the study has advanced the field of radiomics by specifically identifying an exigent unmet need. Overly optimistic claims concerning robustness and generalizability diminish scientific and clinical impact and should be avoided. Publications should extensively report study-design, protocols, detailed quality assurance processes, and standard operating procedures. Although the minute technical details of radiomics are tedious, they can greatly influence robustness, generalizability, and confound meta-analyses. Rigorous reporting guidelines are necessary for radiomics to mature[42–44]. Many journals now encourage and facilitate extensive supplementary materials.

The criteria of the RQS

Overwhelming evidence shows that the quality of reporting of prediction model studies is currently poor[32]. Full and clear reporting of information is required on all aspects of a prediction model to minimize bias and enhance the usefulness of prediction models. An excellent example is the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) initiative[32]. Within this initiative, a set of recommendations was established for the reporting of studies developing, validating, or updating a prediction model, regardless of whether the model serves diagnostic or prognostic purposes. We have emulated this approach in a radiomics-specific context, and suggest that studies should be assessed via the RQS (available online[41]), for which we identified sixteen key components; each assigned a number of points corresponding to the importance of the respective component detailed in TABLE 1.

Translational potential of radiomics

Since the beginning of this decade, radiomics research has advanced dramatically, revealing the potential of this discipline to substantially improve clinical care (TABLE 2). Advances in hardware and software have enabled the realization of clinically feasible quantitative imaging of tissue pathophysiology.

Radiogenomics

Radiogenomics is associated with two closely related but distinct scientific questions: one is the study of the link between germline genotypic variations and the large clinical variability observed in response to radiation therapy[45], and the other is the study of the link between specific imaging traits and specific gene-expression patterns that inform the underlying cellular pathophysiology[46]. Within the radiobiology community, a common hypothesis is that a proportion of the variance in the phenotype of interest (for example, radiation toxicity) can be attributed to genotypic variation. In the clinical context, this hypothesis materializes in a risk of severe treatment-related toxicity for a minority of patients that limits the prescription of potentially curative doses to a majority of patients. We posit that the overall goal of radiogenomics is to isolate the alleles and corresponding radiomic features that underlie the inherited differences in phenotype. Gene-expression profiling of various human tissues has enriched our understanding of cellular pathways and numerous pathological conditions. Investigation of different cancerous tissues in relation to samples of nonmalignant healthy tissue has elucidated tumorigenic processes and assisted in enhanced staging and sub-classification of various malignancies. Gene-expression signatures, each comprised of dozens to hundreds of genes, can meaningfully improve diagnosis, prognosis, and prediction of response to treatment[47–52]. Seminal radiogenomic investigations highlighted the link between radiomic features and gene-expression patterns in patients with cancer[53–55]. One study leveraging survival data in publicly available gene-expression datasets for patients with non-small-cell lung cancer enabled the identification of prognostic imaging biomarkers[54]. This radiogenomics strategy for identifying imaging biomarkers might enable a more-rapid evaluation of novel imaging modalities, thereby accelerating their incorporation into personalized medicine approaches. Another investigation compared clinician-defined features extracted from contrast-enhanced CT images in patients with hepatocellular carcinoma with gene-expression patterns using machine learning with a neural network[55]. Reported combinations of 28 features could reconstruct 78% of the global gene-expression profiles associated with cell proliferation, liver synthetic function, and prognosis. In another study[53], features extracted from MR images to predict global gene-expression patterns in patients with glioblastoma multiforme revealed that an infiltrative phenotype was associated with significantly reduced survival.

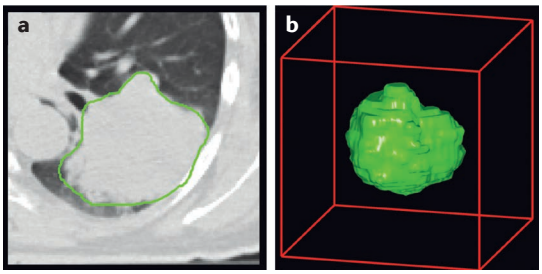

Figure 3 | **Radiomics digital phantom data.** **a** | Representative image of a digital phantom CT image, with the tumour delineation shown outlined in green. **b** | A 3D rendering of the tumour region. The reference feature data is provided (please see the supplementary material).

Table 1 | **The radiomics quality score: RQS**

Criteria		Points
1	Image protocol quality - well-documented image protocols (for example, contrast, slice thickness, energy, etc.) and/or usage of public image protocols allow reproducibility/replicability	+ 1 (if protocols are well-documented) + 1 (if public protocol is used)
2	Multiple segmentations - possible actions are: segmentation by different physicians/algorithms/software, perturbing segmentations by (random) noise, segmentation at different breathing cycles. Analyse feature robustness to segmentation variabilities	+ 1
3	Phantom study on all scanners - detect inter-scanner differences and vendor-dependent features. Analyse feature robustness to these sources of variability	+ 1
4	Imaging at multiple time points - collect images of individuals at additional time points. Analyse feature robustness to temporal variabilities (for example, organ movement, organ expansion/ shrinkage)	+ 1
5	Feature reduction or adjustment for multiple testing - decreases the risk of overfitting. Overfitting is inevitable if the number of features exceeds the number of samples. Consider feature robustness when selecting features	- 3 (if neither measure is implemented) + 3 (if either measure is implemented)
6	Multivariable analysis with non radiomics features (for example, EGFR mutation) - is expected to provide a more holistic model. Permits correlating/inferencing between radiomics and non radiomics features	+ 1
7	Detect and discuss biological correlates - demonstration of phenotypic differences (possibly associated with underlying gene–protein expression patterns) deepens understanding of radiomics and biology	+ 1
8	Cut-off analyses - determine risk groups by either the median, a previously published cut-off or report a continuous risk variable. Reduces the risk of reporting overly optimistic results	+ 1
9	Discrimination statistics - report discrimination statistics (for example, C-statistic, ROC curve, AUC) and their statistical significance (for example, p-values, confidence intervals). One can also apply resampling method (for example, bootstrapping, cross-validation)	+ 1 (if a discrimination statistic and its statistical significance are reported) + 1 (if a resampling method technique is also applied)
10	Calibration statistics - report calibration statistics (for example, Calibration-in-the-large/slope, calibration plots) and their statistical significance (for example, *P*-values, confidence intervals). One can also apply resampling method (for example, bootstrapping, cross-validation)	+ 1 (if a calibration statistic and its statistical significance are reported) + 1 (if a resampling method technique is also applied)
11	Prospective study registered in a trial database - provides the highest level of evidence supporting the clinical validity and usefulness of the radiomics biomarker	+ 7 (for prospective validation of a radiomics signature in an appropriate trial)
12	Validation - the validation is performed without retraining and without adaptation of the cut-off value, provides crucial information with regard to credible clinical performance	- 5 (if validation is missing) + 2 (if validation is based on a dataset from the same institute) + 3 (if validation is based on a dataset from another institute) + 4 (if validation is based on two datasets from two distinct institutes) + 4 (if the study validates a previously published signature) + 5 (if validation is based on three or more datasets from distinct institutes) *Datasets should be of comparable size and should have at least 10 events per model feature
13	Comparison to 'gold standard' - assess the extent to which the model agrees with/is superior to the current 'gold standard' method (for example, TNM-staging for survival prediction). This comparison shows the added value of radiomics	+ 2
14	Potential clinical utility - report on the current and potential application of the model in a clinical setting (for example, decision curve analysis).	+ 2
15	Cost-effectiveness analysis - report on the cost-effectiveness of the clinical application (for example, QALYs generated)	+ 1
16	Open science and data - make code and data publicly available. Open science facilitates knowledge transfer and reproducibility of the study	+ 1 (if scans are open source) + 1 (if region of interest segmentations are open source) + 1 (if code is open source) + 1 (if radiomics features are calculated on a set of representative ROIs and the calculated features and representative ROIs are open source)
	Total points (36 = 100%)	

Table 2 | **Radiomics in practice**

Utility	Modality	Features	Cancer	#Pts	Result	Conclusion	Ref
Tumour prognosis	CT	Intensity, shape, texture, and wavelet	Lung and head & neck	1,019	Lung: (C-index = 0.65, $P = 2.9 \times 10^{-09}$, Wilcoxon test), and a high performance in H&N1 (C-index = 0.69, $P = 8.0 \times 10^{-07}$, Wilcoxon test) and H&N2 (C-index = 0.69, $P = 3.5 \times 10^{-06}$, Wilcoxon test).	Could predict survival in two entirely independent external cohorts of patients, outperforming the current gold standard of TNM status (radiation or concurrent chemoradiation)	1
Tumour prognosis	PET	Texture and shape	Oesophageal	217	A clinical prediction model (C-index = 0.67) was improved by adding radiomic features (C-index = 0.77); however, at a decision threshold of ≥0.9 there was no clear incremental value	Demonstrated that a radiomic PET signature provided statistical incremental value for predicting pathological complete response after preoperative chemoradiation	118
Tumour prognosis	CT	Intensity, texture, and Laplacian of Gaussian filters	Colorectal	326	Training: showed good discrimination (C-index = 0.74) and calibration. Validation: showed good discrimination (C-index = 0.78) and good calibration	Decision curve analysis demonstrated that a final nomogram consisting of the radiomic portal venous-phase CT signature, CT-reported lymph-node status, and carcinoembryonic antigen level was clinically useful	119
Distant metastasis	CT	Texture, Laplacian of Gaussian and wavelet filters	Lung	182	Could predict distant metastasis in an independent validation dataset (C-index = 0.61, $P = 1.79 \times 10^{-17}$). Adding this radiomic-signature to a clinical model resulted in a significant improvement ($P = 1.56 \times 10^{-11}$)	Provided superior information than clinical data capturing detailed information of the tumour phenotype and can be used as a prognostic biomarker for distant metastasis	120
Distant metastasis	CT	Wavelet filters	Lung	113	Significantly prognostic for distant metastasis (C-index = 0.67, q-value <0.1), while none of the conventional and clinical parameters were prognostic. Three conventional and four radiomic features were prognostic for overall survival	Demonstrates that radiomic features have potential to be prognostic for some outcomes that conventional imaging metrics cannot predict in patients receiving stereotactic body radiation therapy	121
Efficacy	CT	Intensity and texture	Oesophageal	106	Significant change in radiomics feature values was observed with increasing radiation dose (pre and post radiotherapy scans). AUC = 0.75 using multiple features in a classifier	Demonstrated the ability to individualize the measurement of patient lung tissue reaction to radiotherapy and assess radiation pneumonitis development	122
Staging	CT	Intensity, texture, Laplacian of Gaussian filters	Colorectal	494	Training: AUC = 0.79, $P = <0.0001$. Validation: AUC = 0.71, $P = <0.0001$. The radiomics signature was an independent predictor for staging	Demonstrated the ability to discriminate between stage I–II from III–IV, which may serve as a complementary tool for the preoperative tumour staging	123
Screening	CT	Intensity and texture	Lung	196	AUC = 0.83, $P = <0.05$. Radiomics performance was commensurate with the McWilliams risk-assessment model	Demonstrated that radiomics at baseline can be used to assess risk for the development of cancer	124
Survival	MR	Volumetric	Brain	141	C-Index = 0.60, $P = 4 \times 10^{-4}$. Volumetric features were significantly associated with diverse sets of biological processes, FDR <0.05	Demonstrated the ability to derive the biological state of a glioblastoma tumour that can be used to develop personalized treatment strategies	125
Survival	CT	Texture	Lung	282	C-Index = 0.72, improved accuracy of calibration and the classification of survival outcomes (net reclassification improvement: 0.182, $P = 0.02$).	Decision curve analysis demonstrated that in terms of clinical usefulness, the radiomics nomogram outperformed the traditional staging system and the clinical-pathologic nomogram	126
Tumour prognosis	CT	Intensity, shape, texture, and wavelet	Oropharyngeal	542	C-Index = 0.63, $P = 2.72 \times 10^{-9}$. Kaplan-Meier survival curves were significantly different ($P < 0.05$) between high and low radiomic signature model predictions for all cohorts	Demonstrated external validation of the signature, the signature had significant prognostic power irrespective of the presence or absence of CT artifacts	127

Table 2 (cont.) | **Radiomics in practice**

Utility	Modality	Features	Cancer	#Pts	Result	Conclusion	Ref
Overall survival	PET	Shape, intensity, and texture	Pancreatic	139	C-Index = 0.66, significantly better than competing prognostic indices (0.48–0.64, Wilcoxon rank sum test $P = 1 \times 10^{-6}$)	Demonstrated external validation of the signature if validated in large, prospective cohorts, the signature might be used to identify patients for individualized risk-adaptive therapy	128
Recurrence	MR	Shape, intensity, and texture	Breast	89	AUC = 0.88, 0.76, and 0.68 for MammaPrint, Oncotype DX, and PAM50 risk of relapse based on subtype respectively, all statistically significant, $P = \leq 0.05$	Demonstrates that breast MR imaging radiomics shows promise for image-based phenotyping in assessing the risk of breast cancer recurrence	129

Nevertheless, gene-expression profiling relies on surgical procurement of sampled tissue specimens associated with multiple risks and potential complications, consequently rendering it unfeasible for many patients with cancer. In stark contrast to genetic profiling studies, radiomic features[53,55–59] capture intratumoural heterogeneity in a non-invasive three-dimensional manner, and can be obtained as part of routine clinical care. For example, approximately 15% of triple-negative breast cancers diagnosed globally are associated with poor outcomes after treatment[60]. Reliable techniques for the assessment of HER2 expression (typically fluorescence *in situ* hybridization (FISH)) are expensive and time-consuming. In a study that considered the tumour as well as its surrounding parenchyma, DCE–MRI radiomic image phenotyping provided useful information for the diagnosis of triple-negative breast tumours[61]. Currently, the radiogenomics landscape is evolving rapidly from an effort to screen a limited number of candidate genes towards an open-discovery approach into the powerful, but challenging, era of RLHC[62–66].

Radiogenomic features provide valuable biomarkers for CDSS[67–70]; these include prognostic and predictive factors for outcomes[71], such as tumour response to treatment as well as nonmalignant-tissue tolerance to the same treatment. Notwithstanding these virtues, trials of radiogenomics biomarkers are susceptible to experimental and imaging inconsistency; therefore, standardization of assay criteria, image acquisition, segmentation, trial design, and an analytical approach is vital if radiogenomics biomarkers are to be effective diagnostic, prognostic, and predictive tools in oncology[72].

Radiosensitivity and the tumour habitat

Tumour control following radiotherapy is governed by the following criteria: the quantity of cancer stem cells present in the tumour (which is characteristically associated with pretreatment tumour volume), the innate radiosensitivity of the stem cells, the hypoxic fraction, reoxygenation of the tumour vicinity and/or repopulation capacity throughout the course of therapy[73–75].

Radiogenomic analysis of tumour microenvironments has the capacity to unlock knowledge with respect to the aforementioned criteria[76,77]. An example is the strong correlation between microvascular density and PET–MR-derived radiomics features reported in patients with primary clear-cell-renal-cell-carcinoma[78]. In general, tumours display considerable variability in radiosensitivity, which even affects tumour cells of analogous origin or histological type[79–81]. The quantification of the radiosensitivity of human tumours is presently performed on the basis of the *ex vivo* tumour survival fraction, and the detection of unrepaired DNA double-strand breaks[82,83]. Preclinical (including prostate, lung, and brain cancers) and clinical (such as cervical, and head-and-neck cancers) studies have proven that tumour-cell radiosensitivity is a key feature for the prediction of outcomes to radiotherapy in patients with prostate, lung, brain, cervical, or head-and-neck cancers[84–88]. These data, however, were built on results from colony assays that involved technical deficiencies, such as poor plating efficiency (<70%) for human tumours or protracted time required to produce data (up to several weeks).

Overall, these approaches have been undermined by the presence of substantial experimental variability rather than by the existence of interpatient variations in radiosensitivity. Non-malignant tissue toxicity is the dose-limiting factor in radiation oncology; therefore, a comprehensive CDSS should be built upon predictors of dose, tumour-control versus non-malignant-tissue-complication probability ratio, as well as cost-effectiveness[89], in order to facilitate improved escalated or de-escalated individualized treatments for patients.

Immunotherapy

In the field of oncology, a promising research area is that of biomarkers — in particular, biomarkers for immunotherapy and imaging biomarkers[90,91]. The stimulation of an antitumour immune response in response to radiotherapy has been well documented[92]. For a robust and vigorous immune response to be elicited, the activation of antigen-specific T-cells coupled with memory effects is required[93]. Such an immune response is dependent on the expression of specific antigens on tumour cells (neoantigens), which are subsequently identified by the immune system. Neoantigens arise from mutated proteins within the tumour cell. Research results demonstrate that the success rate of immunotherapies relies on the presence of neoantigen-specific T-cells[94]. Moreover, the mutational load of many human tumour types correlates with the cytolytic activity of natural killer cells and T-cells[95]. When a tumour has the potential to be identified by the immune system, a suitable immune response can be coordinated. Importantly, neoantigens must be taken up by antigen-presenting cells (APCs) or dendritic cells (DCs) and subsequently

cross-presented to naive T-cells[96]. T cells are then converted into tumour-killing cytotoxic T cells. The capacity of radiation to boost immune responses seems to be crucially dependent on the quantity and quality of DCs present in the tumour local environment[97,98].

Tumours have diverse means to shield from ample cytotoxic T-cell responses; for example, by interfering with several immune checkpoints, such as PD-L1, a protein with biomarker potential. The mutational load of the tumour has been correlated with clinical responses to anti-PD-1-mediated immunotherapy, for example, in patients with non-small-cell lung cancer[99,100]. Such biological and genetic features hold potential; radiogenomic analysis will undoubtedly have an important role in leveraging this information in future CDSS. The basic hypothesis, still to be tested, is that tumours with high mutational loads have more neoantigens and, consequently, will be more heterogeneous on radiomic analysis and more sensitive to immunotherapy than tumours with low mutational loads. The exact opposite scenario is noted in response to radiotherapy or chemotherapy; in these situations enhanced radiomics-assessed tumour heterogeneity is an adverse prognostic factor. Regarding the combination of radiotherapy and/or chemotherapy and immunotherapy, no clear indication exists as to which effect will dominate.

Technical aspects

Accredited radiogenomic centres should be established. Stakeholders in the academic, clinical, industry, and regulatory spheres must collaborate to create, sustain, and standardize the required best-practice framework. Radiomic studies are difficult to perform consistently, and thus, accreditation is vital to the advancement of radiomics. Techniques for the workflow of radiomics ought to be independent of vendors and upgrades to hardware and/or software. Radiomic studies should incorporate reproducibility assessments owing to the beneficial ethical, economic and logistical effects they have (such as informing power calculations and required samples sizes, multicentric trial duration and trial cost). Optimal reproducibility and stability enables multicentre studies to maximize the likelihood of a validated radiomic signature being fit-for-purpose in routine clinical use. Prospective studies sufficiently powered to relate radiomics data to clinical outcomes in appropriate patient populations are pivotal. Indeed, numerous studies are underpowered for sensitivity and specificity; however, study populations should not be skewed by selecting only those patients who are more capable of adhering to complex imaging protocols than the general population. All findings should be published, including true-negatives, false-negatives and false-positives, and the perceived adversity to negative results tempered because the inclusion of substantial bias risks distorting the radiomics landscape.

Economic elements

Multicentric, collaborative and federated efforts are required to share, store, and curate data. Data-sharing enables initiation of highly powered prospective studies and accelerates the development and validation of radiomic signatures derived from new and existing data. Trials conducted in centres that are part of a network can quickly recruit sufficient patient numbers to drive discovery and innovation. Outcome studies should include health economic considerations. Moreover, cost-per-quality-adjusted-life-year comparisons should be conducted with and without radiomics to more accurately determine the economic potential of such studies[101].

The way forward for radiomics

Virtual biopsy

In patients with cancer, different parts of the tumour have distinct molecular characteristics; such differences change over time. As it is not possible to biopsy every part of each tumour at multiple time points, the optimal characterization of tumours is not achieved using biopsy samples (FIG. 4).

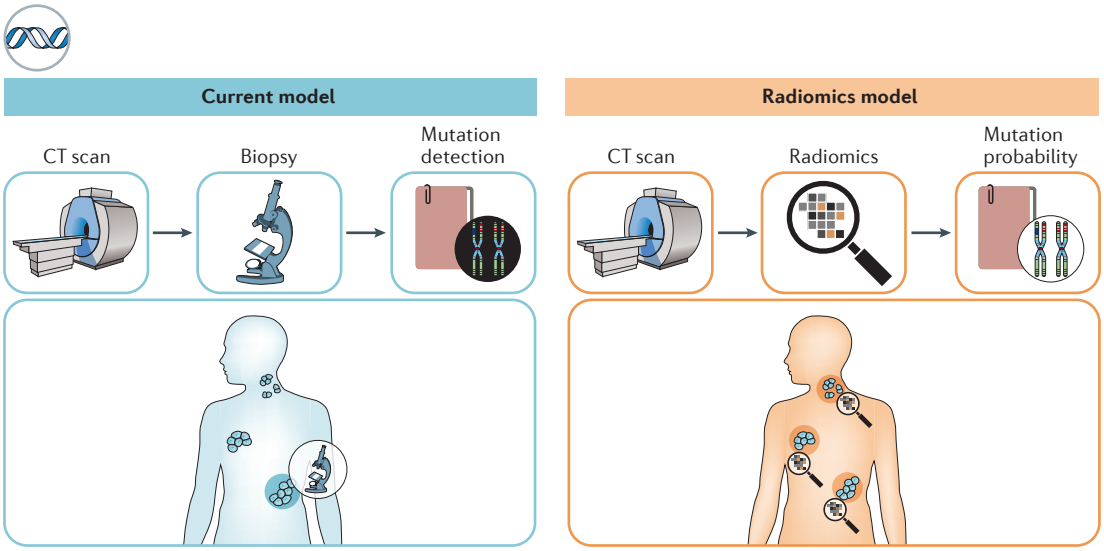

Figure 4 | **Radiogenomics analysis can reveal relationships between imaging phenotypes and gene-expression patterns.** Such relationships can include expressions of individual genes as well as measures that summarize expressions of specific gene subsets.

Delta radiomics

Published work mainly focused on imaging data acquired at a single time point, mostly imaging tumours before the start of treatment. Delta-radiomics introduces a time component and comprises extraction of quantitative features from image sets acquired over the course of treatment[102–104], which provides information on the evolution of feature values[105] (FIG. 5). Delta-radiomics promises to improve diagnosis, prognosis, prediction, monitoring, image-based intervention, or assessment of therapeutic response.

Infrastructure for radiomics

Radiomics demonstrates huge potential to deepen knowledge and broaden the horizons of imaging, in order to achieve greater precision and extraction of *in vivo* biological information. To fully harness the potential of radiomics, research and clinical communities must embrace an interdisciplinary shared vision of precision medicine. Extracted radiomic features must be stored in searchable databases in order to realize the unprecedented potential for RLHC that routine standard-of-care imaging represents. Hence, RLHC networks can dynamically capture multimodal data and share knowledge across departmental and institutional boundaries, in order to accumulate sufficient datasets for significant statistical power in model development and validation.

Big data

Ideal RLHCs necessitate the 4Vs of 'big data': volume, variety, velocity, and veracity of data. The volume of data is important: firstly, the quality of knowledge gained from a study is correlated with the number of patients

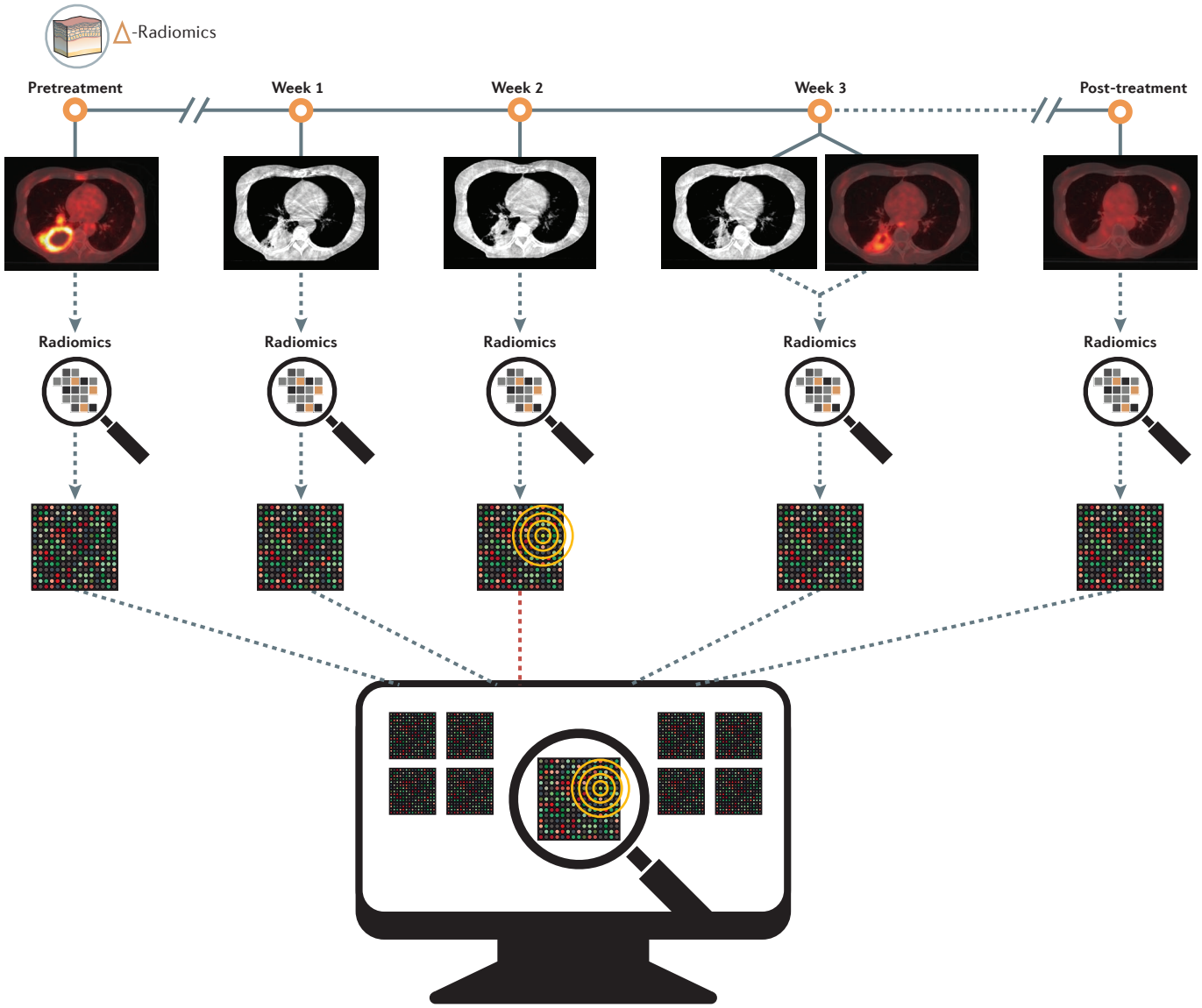

Figure 5 | **Schematic overview of a clinical decision-support system graphical user interface illustrating the concept of delta-radiomics.** In this example, a clinician requests the radiomic analysis of a patient on the basis of combined longitudinal PET–CT images, potentially enabling improved diagnosis, early response prediction, improved clinical decision-making and, consequently, a better prognosis.

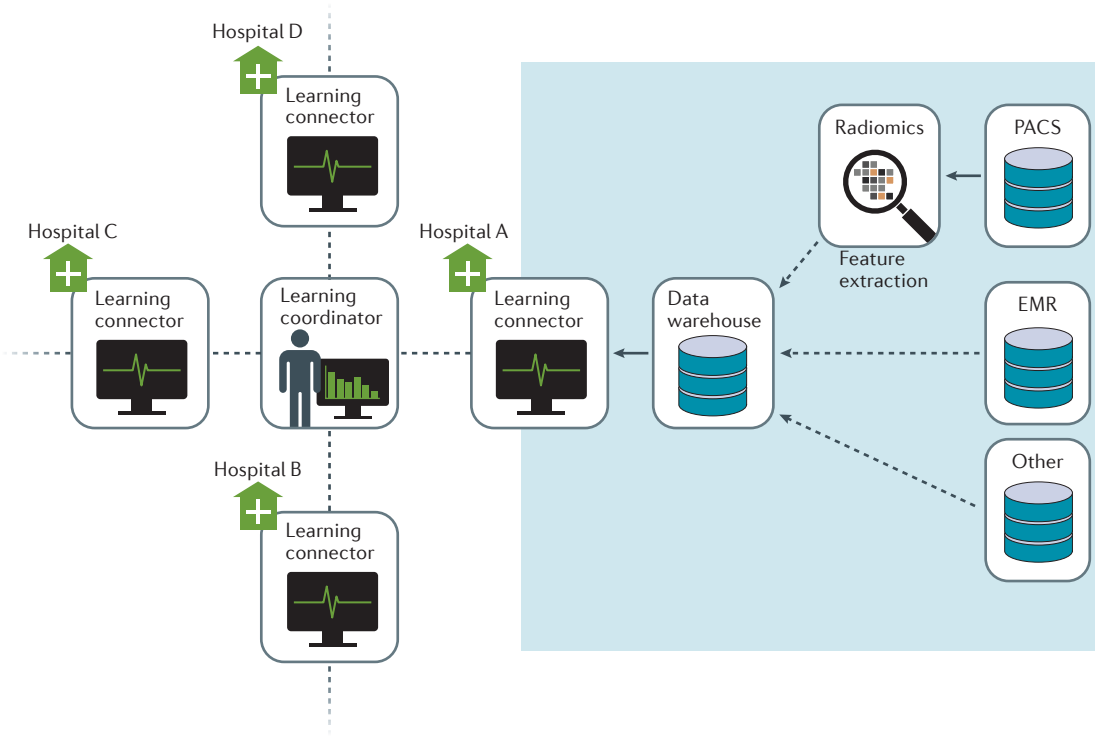

Figure 6 | **Schematic diagram of the CAT system.** Multiple centres are linked via learning connectors. The connector is the interface where machine learning algorithms (sent from the learning coordinator) learn models from local data. Of note, privacy-sensitive information remains in the institute. Partner sites exist in the Netherlands, Germany, Belgium, Italy, Denmark, Australia, China, India, South Africa, Ireland, UK, USA and Canada. The system is built from a combination of open-source information communication technologies and can deliver data locally via SQL query, or to the wider CAT network via a SPARQL end point. CAT; Computer-assisted theragnostic EMR; electronic medical record; PACS, picture archiving and communication system; SPARQL; simple protocol and RDF query language.

from whom data were obtained. Secondly, a greater data volume enables the acquisition of more variables in the model development phase. Thirdly, knowledge related to patients with rare disease variants is more easily gained from larger datasets. The variety of data, both in terms of treatment and of patient characteristics, is critical for deciding which treatment is optimal for each individual patient. The velocity of data acquisition is important to guarantee that knowledge is gathered as swiftly and perpetually as possible, while the veracity of data is critical to the amount of confidence that can be ascribed to the knowledge gained.

Data sharing

Procuring data of sufficient quality with regard to the 4Vs is central to RLHCs. A pressing need to embrace knowledge and data-sharing technology[106], which transcends institutional and national boundaries[107], drives both the research and clinical communities. The following established obstacles to data sharing[108] are apparent in the medical domain: insufficient human resources or insufficient time, cultural and language difficulties, data recording methods, the political and academic value of data, hazards to reputation, or legal and privacy considerations, to name a few. These issues, although not easy to overcome, must be addressed.

One initiative to accomplish this goal is CancerLinQ[109], the ASCO data centralization approach. Another initiative is worldCAT that consists of a novel data-federated approach that successfully links radiotherapy institutes in the Netherlands, Germany, Belgium, Italy, Denmark, Australia, China, India, South Africa, Ireland, UK, USA and Canada (FIG. 6)[110]. In addition, universal streamlined solutions through advanced information communication technologies have been central to the realization of this endeavor, readily facilitating synchronized RLHC in each centre without inclusion of sensitive data, which overcomes the classic barriers to data sharing. Other important links include The Cancer Imaging Archive (TCIA)[111], The Quantitative Imaging Network (QIN)[112], the Quantitative Imaging Biomarkers Alliance (QIBA)[113], and Quantitative Imaging in Cancer: Connecting Cellular Processes with Therapy (QuIC-ConCePT)[114].

Ontologies for learning

For RLHCs to succeed, the creation of data with semantic interoperability, also known as 'machine-readable' data[115] is needed, in which local terms are harmonized from concepts of well-defined ontologies (such as the NCI Thesaurus or ICD-10). Exploiting this technique, the ontology terms serve as a common reference for the data at each institutional site, permitting a unified process for information retrieval enabled by a semantic gateway to the data. A benefit of this approach is that it promotes standardization with respect to data management (such as disease-specific 'umbrella' protocols: NCT01855191)[116,117].

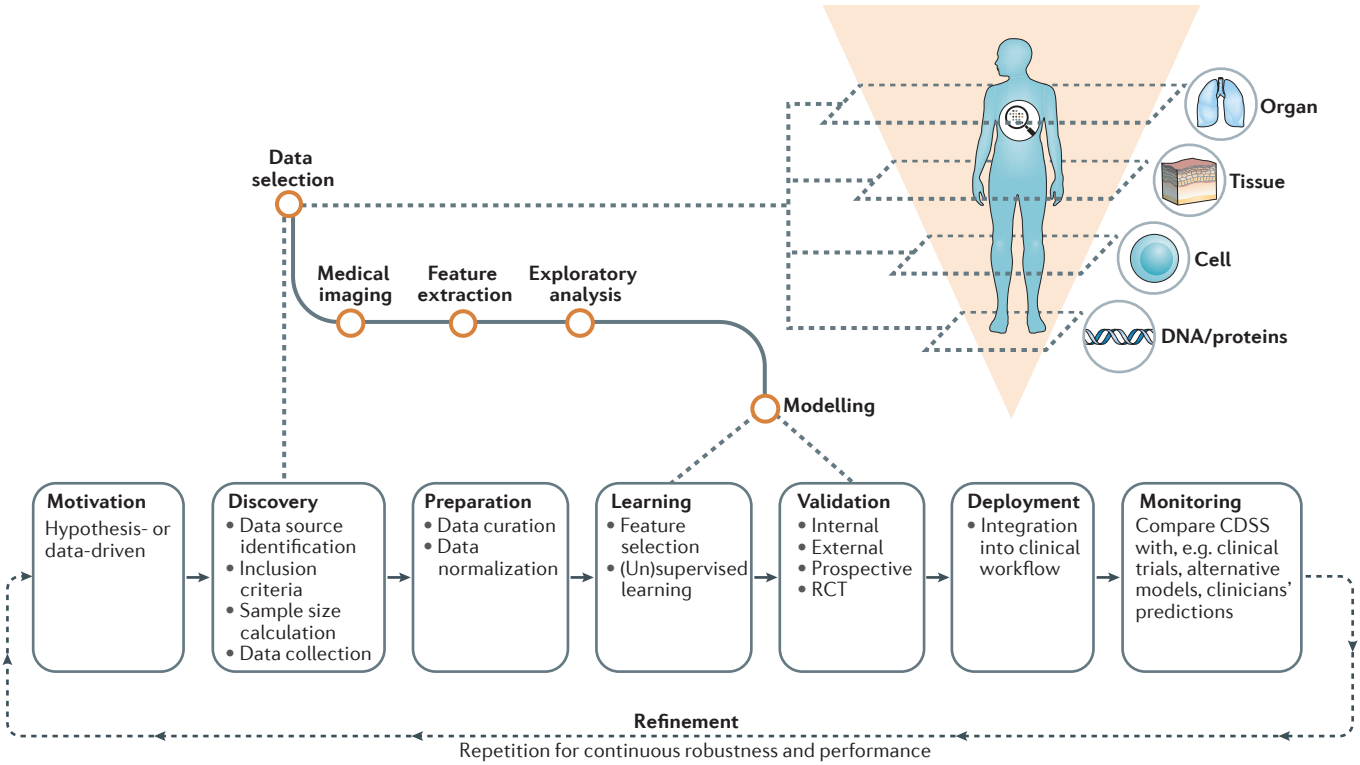

Figure 7 | **Overview of the methodological processes for RLHC and how the radiomics workflow fits into the development of a CDSS.** Data selection, discovery, collection and preparation, model(s) development/validation and implementation, assessment of clinical utility and ultimately refinement through continuous repetition of the process (quality control and assurance protocols are requisite throughout the process).

Role of radiomics in the future

Picture archiving and radiomics knowledge systems (PARKS) of the future will identify, segment, and extract features from regions of interest. If previous images associated with the same patient are accessible, the earlier identified regions of interest will be automatically identified by the PARKS software. Quantitative image features that are uploaded to a shared database and compared with previous images will be automatically extracted by the PARKS to enhance CDSS for diagnosis, prognosis, and treatment, resulting in improved personalization and precision medicine (FIG. 7). Such capabilities are on the technological, scientific, and clinical horizons, as most current picture archiving and communication systems have the capability to co-register current images with previous images and perform user-interactive segmentation. For the immediate future, the field of radiomics will focus on the creation of suitable infrastructures for powerful RLHC networks that will facilitate the development and validation of models.

Conclusions

Our vision for radiomics is expansive and bold. In the reasonably near future, we envision that CDSS that apply knowledge leveraged from radiomic features mined from global RLHC networks populated by standard-of-care imaging will enable increased personalized delivery of medicine. For this vision to be updated within the routine clinical setting, clinicians and medical physicists must be incentivized to participate in the process. Moreover, standardization is crucial to this endeavor, principally in the acquisition of high-quality data. Standardization underpins coherent clinical guidelines with agreed standards for image acquisition and analysis, as well as data-sharing techniques that exploit matching ontologies. Continuous re-evaluation and demonstration of the clinical utility of a CDSS is as significant as standardizing the development and validation of the design of clinical trials. These crucial steps are the foundation of a successful CDSS. Simultaneous and synergistic advances in RLHC and radiomics will empower the next major breakthroughs in personalization and precision medicine.

1. Aerts, H. *et al.* Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **5**, 4006 (2014).
2. Hood, L. & Friend, S. H. Predictive, personalized, preventive, participatory (P4) cancer medicine. *Nat. Rev. Clin. Oncol.* **8**, 184–187 (2011).
3. Lambin, P. *et al.* Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **48**, 441–446 (2012).
4. Kumar, V. *et al.* Radiomics: the process and the challenges. *Magn. Reson. Imaging* **30**, 1234–1248 (2012).
5. Haase, A. T. *et al.* Quantitative image analysis of HIV-1 infection in lymphoid tissue. *Science* **274**, 985–989 (1996).
6. Lambin, P. *et al.* Predicting outcomes in radiation oncology — multifactorial decision support systems. *Nat. Rev. Clin. Oncol.* **10**, 27–40 (2013).
7. [No authors listed] Medicine: Computers by the Bedside. *Nature* **224**, 636–637 (1969).
8. Schoolman, H. & Bernstein, L. Computer use in diagnosis, prognosis, and therapy. *Science* **200**, 926–931 (1978).
9. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: images are more than pictures, they are data. *Radiology* **278**, 563–577 (2016).
10. Roelofs, E. *et al.* International data-sharing for radiotherapy research: an open-source based infrastructure for multicentric clinical data mining. *Radiother. Oncol.* **110**, 370–374 (2014).

11. Roelofs, E. *et al.* Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial. *Radiother. Oncol.* **108**, 174–179 (2013).
12. Miotto, R., Li, L., Kidd, B. A. & Dudley, J. T. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci. Rep.* **6**, 26094 (2016).
13. Nead, K. T. *et al.* Androgen deprivation therapy and future alzheimer's disease risk. *J. Clin. Oncol.* **34**, 566–571 (2016).
14. Gatenby, R. A., Grove, O. & Gillies, R. J. Quantitative imaging in cancer evolution and ecology. *Radiology* **269**, 8–14 (2013).
15. Aerts, H. L. The potential of radiomic-based phenotyping in precision medicine: A review. *JAMA Oncol.* **2**, 1636–1642 (2016).
16. Lambin, P. *et al.* Decision support systems for personalized and participative radiation oncology. *Adv. Drug Delivery Rev.* **109**, 131–153 (2017).
17. Vickers, A. Prediction models: revolutionary in principle, but do they do more good than harm? *J. Clin. Oncol.* **29**, 2951–2952 (2011).
18. Yip, S. S. & Aerts, H. J. Applications and limitations of radiomics. *Phys. Med. Biol.* **61**, R150–166 (2016).
19. Polan, D. F., Brady, S. L. & Kaufman, R. A. Tissue segmentation of computed tomography images using a Random Forest algorithm: a feasibility study. *Phys. Med. Biol.* **61**, 6553–6569 (2016).
20. Balagurunathan, Y. *et al.* Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol.* **7**, 72–87 (2014).
21. Grootjans, W. *et al.* The impact of optimal respiratory gating and image noise on evaluation of intra-tumor heterogeneity in ^{18}F-FDG PET imaging of lung cancer. *J. Nucl. Med.* **57**, 1692–1698 (2016).
22. Larue, R. T., Defraene, G., de Ruysscher, D., Lambin, P. & van Elmpt, W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br. J. Radiol.* **90**, 20160665 (2017).
23. Mackin, D. *et al.* Measuring computed tomography scanner variability of radiomics features. *Invest. Radiol* **50**, 757–765 (2015).
24. Balagurunathan, Y. *et al.* Test–retest reproducibility analysis of lung CT image features. *J. Digit. Imaging* **27**, 805–823 (2014).
25. Zhao, B. *et al.* Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non–small cell lung cancer. *Radiology* **252**, 263–272 (2009).
26. Zhao, B. *et al.* Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci. Rep.* **6**, 23428 (2016).
27. Hatt, M. *et al.* Characterization of PET/CT images using texture analysis: the past, the present… any future? *Eur. J. Nucl. Med. Mol. Imaging* **44**, 151–165 (2016).
28. Fang, Y. H. *et al.* Development and evaluation of an open-source software package "CGITA" for quantifying tumor heterogeneity with molecular images. *Biomed. Res. Int.* **2014**, 248505 (2014).
29. Zhang, L. *et al.* IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. *Med. Phys.* **42**, 1341–1353 (2015).
30. Parmar, C., Grossmann, P., Bussink, J., Lambin, P. & Aerts, H. J. Machine learning methods for quantitative radiomic biomarkers. *Sci. Rep.* **5**, 13087 (2015).
31. https://github.com/ (2017 May 18th).
32. Collins, G., Reitsma, J., Altman, D. & Moons, K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *Ann. Intern. Med.* **162**, 55–63 (2015).
33. Lemeshow, S. & Hosmer, D. W. Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am. J. Epidemiol.* **115**, 92–106 (1982).
34. Debray, T. P. *et al.* A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J. Clin. Epidemiol.* **68**, 279–289 (2015).
35. Steyerberg, E. *et al.* Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* **21**, 128–138 (2010).
36. Leek, J. T. & Peng, R. D. Statistics: P values are just the tip of the iceberg. *Nature* **520**, 612 (2015).
37. Drummond, C. Replicability is not reproducibility: nor is it good science. In *Evaluation Methods for Machine Learning* (2009).
38. Peng, R. D. Reproducible research in computational science. *Science* **334**, 1226–1227 (2011).
39. Peng, R. D., Dominici, F. & Zeger, S. L. Reproducible epidemiologic research. *Am. J. Epidemiol.* **163**, 783–789 (2006).
40. Lambin, P. Radiomics digital phantom. *CancerData. org* https://www.cancerdata.org/resource/doi%3A10.17195/candat.2016.08.1 (2017).
41. http://www.radiomics.world/ (2017 May 18th).
42. Altman, D. G., McShane, L. M., Sauerbrei, W. & Taube, S. E. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *BMC Med.* **10**, 51 (2012).
43. Pepe, M. S. & Feng, Z. Improving biomarker identification with better designs and reporting. *Clin. Chem.* **57**, 1093–1095 (2011).
44. Poste, G. Biospecimens, biomarkers, and burgeoning data: the imperative for more rigorous research standards. *Trends Mol. Med.* **18**, 717–722 (2012).
45. Rosenstein, B. S. *et al.* Radiogenomics: radiobiology enters the era of big data and team science. *Int. J. Radiat. Oncol. Biol. Phys.* **89**, 709–713 (2014).
46. Rutman, A. M. & Kuo, M. D. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. *Eur. J. Radiol.* **70**, 232–241 (2009).
47. Chang, H. Y. *et al.* Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc. Natl Acad. Sci. USA* **102**, 3738–3743 (2005).
48. Chen, X. *et al.* Gene expression patterns in human liver cancers. *Mol. Biol. Cell* **13**, 1929–1939 (2002).
49. Chung, C. H., Bernard, P. S. & Perou, C. M. Molecular portraits and the family tree of cancer. *Nat. Genet.* **32** (Suppl.), 533–540 (2002).
50. Paik, S. *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351**, 2817–2826 (2004).
51. Paik, S. *et al.* Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J. Clin. Oncol.* **24**, 3726–3734 (2006).
52. Segal, E., Friedman, N., Kaminski, N., Regev, A. & Koller, D. From signatures to models: understanding cancer using microarrays. *Nat. Genet.* **37**, S38–S45 (2005).
53. Diehn, M. *et al.* Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. *Proc. Natl Acad. Sci. USA* **105**, 5213–5218 (2008).
54. Gevaert, O. *et al.* Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results. *Radiology* **264**, 387–396 (2012).
55. Segal, E. *et al.* Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nat. Biotechnol.* **25**, 675–680 (2007).
56. Gao, X. *et al.* The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from ^{18}F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur. J. Radiol.* **84**, 312–317 (2015).
57. Harry, V. N., Semple, S. I., Parkin, D. E. & Gilbert, F. J. Use of new imaging techniques to predict tumour response to therapy. *Lancet Oncol.* **11**, 92–102 (2010).
58. O'Connor, J. P. *et al.* Quantitative imaging biomarkers in the clinical development of targeted therapeutics: current and future perspectives. *Lancet Oncol.* **9**, 766–776 (2008).
59. Panth, K. M. *et al.* Is there a causal relationship between genetic changes and radiomics-based image features? An *in vivo* preclinical experiment with doxycycline inducible GADD34 tumor cells. *Radiother. Oncol.* (2015).
60. Jemal, A. *et al.* Global cancer statistics. *CA Cancer J. Clin.* **61**, 69–90 (2011).
61. Wang, J. *et al.* Identifying triple-negative breast cancer using background parenchymal enhancement heterogeneity on dynamic contrast-enhanced MRI: a pilot radiomics study. *PLoS ONE* **10**, e0143308 (2015).
62. Abernethy, A. *et al.* Rapid-learning system for cancer care. *J. Clin. Oncol.* **28**, 4268–4274 (2010).
63. Lambin, P. *et al.* Modern clinical research: How rapid learning health care and cohort multiple randomised clinical trials complement traditional evidence based medicine. *Acta Oncol.* **54**, 1289–1300 (2015).
64. Dekker, A. *et al.* Rapid learning in practice: A lung cancer survival decision support system in routine patient care data. *Radiother. Oncol.* **113**, 47–53 (2014).
65. Ginsburg, G., Staples, J. & Abernethy, A. Academic medical centers: ripe for rapid-learning personalized health care. *Sci. Transl. Med.* **3**, 101cm27 (2011).
66. Lambin, P. *et al.* Rapid learning health care in oncology — An approach towards decision support systems enabling customised radiotherapy. *Radiother. Oncol.* **109**, 159–164 (2013).
67. Buettner, R., Wolf, J. & Thomas, R. K. Lessons learned from lung cancer genomics: the emerging concept of individualized diagnostics and treatment. *J. Clin. Oncol.* **31**, 1858–1865 (2013).
68. Colen, R. *et al.* NCI Workshop Report: clinical and computational requirements for correlating imaging phenotypes with genomics signatures. *Transl Oncol.* **7**, 556–569 (2014).
69. Rizzo, S. *et al.* CT radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer. *Eur. Radiol.* **26**, 32–42 (2015).
70. Taguchi, F. *et al.* Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. *J. Natl Cancer Inst.* **99**, 838–846 (2007).
71. Yaromina, A., Krause, M. & Baumann, M. Individualization of cancer treatment from radiotherapy perspective. *Mol. Oncol.* **6**, 211–221 (2012).
72. Dancey, J. E. *et al.* Guidelines for the development and incorporation of biomarker studies in early clinical trials of novel agents. *Clin. Cancer Res.* **16**, 1745–1755 (2010).
73. Krause, M., Yaromina, A., Eicheler, W., Koch, U. & Baumann, M. Cancer stem cells: targets and potential biomarkers for radiotherapy. *Clin. Cancer Res.* **17**, 7224–7229 (2011).
74. Lindegaard, J. C., Overgaard, J., Bentzen, S. M. & Pedersen, D. Is there a radiobiologic basis for improving the treatment of advanced stage cervical cancer? *J. Natl Cancer Inst. Monogr.* **21**, 105–112 (1996).
75. Yaromina, A. *et al.* Pre-treatment number of clonogenic cells and their radiosensitivity are major determinants of local tumour control after fractionated irradiation. *Radiother. Oncol.* **83**, 304–310 (2007).
76. Lambin, P. *et al.* The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. *Radiother. Oncol.* **96**, 145–152 (2010).
77. Prokopiou, S. *et al.* A proliferation saturation index to predict radiation response and personalize radiotherapy fractionation. *Radiat. Oncol.* **10**, 159 (2015).
78. Yin, Q. *et al.* Associations between tumor vascularity, vascular endothelial growth factor expression and PET/MRI radiomic signatures in primary clear-cell-renal-cell-carcinoma: proof-of-concept study. *Sci. Rep.* **7**, 43356 (2017).
79. Menegakis, A. *et al.* Residual γH2AX foci after *ex vivo* irradiation of patient samples with known tumour-type specific differences in radio-responsiveness. *Radiother. Oncol.* **116**, 480–485 (2015).
80. Menegakis, A. *et al.* γH2AX assay in *ex vivo* irradiated tumour specimens: A novel method to determine tumour radiation sensitivity in patient-derived material. *Radiother. Oncol.* **116**, 473–479 (2015).
81. Slonina, D. & Gasinska, A. Intrinsic radiosensitivity of healthy donors and cancer patients as determined by the lymphocyte micronucleus assay. *Int. J. Radiat. Biol.* **72**, 693–701 (1997).
82. Fertil, B. & Malaise, E. P. Intrinsic radiosensitivity of human cell lines is correlated with radioresponsiveness of human tumors: analysis of 101 published survival curves. *Int. J. Radiat. Oncol. Biol. Phys.* **11**, 1699–1707 (1985).
83. Menegakis, A. *et al.* Prediction of clonogenic cell survival curves based on the number of residual DNA double strand breaks measured by γH2AX staining. *Int. J. Radiat. Biol.* **85**, 1032–1041 (2009).
84. Bjork-Eriksson, T., West, C., Karlsson, E. & Mercke, C. Tumor radiosensitivity (SF2) is a prognostic factor for local control in head and neck cancers. *Int. J. Radiat. Oncol. Biol. Phys.* **46**, 13–19 (2000).
85. Chitnis, M. M. *et al.* IGF-1R inhibition enhances radiosensitivity and delays double-strand break repair by both non-homologous end-joining and homologous recombination. *Oncogene* **33**, 5262–5273 (2014).
86. Du, S. *et al.* Attenuation of the DNA damage response by transforming growth factor-beta inhibitors enhances radiation sensitivity of non-small-cell lung cancer cells *in vitro* and *in vivo*. *Int. J. Radiat. Oncol. Biol. Phys.* **91**, 91–99 (2015).

87. Kahn, J. *et al.* The mTORC1/mTORC2 inhibitor AZD2014 enhances the radiosensitivity of glioblastoma stem-like cells. *Neuro Oncol.* **16**, 29–37 (2014).
88. West, C. M., Davidson, S. E., Roberts, S. A. & Hunter, R. D. The independence of intrinsic radiosensitivity as a prognostic factor for patient response to radiotherapy of carcinoma of the cervix. *Br. J. Cancer* **76**, 1184–1190 (1997).
89. Cheng, Q. *et al.* Development and evaluation of an online three-level proton versus photon decision support prototype for head and neck cancer — Comparison of dose, toxicity and cost-effectiveness. *Radiother. Oncol.* **118**, 281–285 (2016).
90. Okada, H. *et al.* Immunotherapy response assessment in neuro-oncology: a report of the RANO working group. *Lancet Oncol.* **16**, e534–e542 (2015).
91. Tang, C. *et al.* Pathology-based non-small cell lung cancer radiomics signature describing the local tumor immune environment: discovery and validation. *Int. J. Radi. Oncol. Biol. Phys.* **96**, S42–S43 (2016).
92. Formenti, S. C. & Demaria, S. Combining radiotherapy and cancer immunotherapy: a paradigm shift. *J. Natl Cancer Inst.* **105**, 256–265 (2013).
93. Coulie, P. G., Van den Eynde, B. J., van der Bruggen, P. & Boon, T. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat. Rev. Cancer* **14**, 135–146 (2014).
94. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* **348**, 69–74 (2015).
95. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).
96. Mellman, I. & Steinman, R. M. Dendritic cells: specialized and regulated antigen processing machines. *Cell* **106**, 255–258 (2001).
97. Demaria, S., Golden, E. B. & Formenti, S. C. Role of local radiation therapy in cancer immunotherapy. *JAMA Oncol.* **1**, 1325–1332 (2015).
98. Golden, E. B. *et al.* Local radiotherapy and granulocyte-macrophage colony-stimulating factor to generate abscopal responses in patients with metastatic solid tumours: a proof-of-principle trial. *Lancet Oncol.* **16**, 795–803 (2015).
99. Garon, E. B. *et al.* Pembrolizumab for the treatment of non-small-cell lung cancer. *N. Engl. J. Med.* **372**, 2018–2028 (2015).
100. Rizvi, N. A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).
101. Sanghera, S., Barton, P., Bhattacharya, S., Horne, A. W. & Roberts, T. E. Pharmaceutical treatments to prevent recurrence of endometriosis following surgery: a model-based economic evaluation. *BMJ Open* **6**, e010580 (2016).
102. Carvalho, S. *et al.* Early variation of FDG-PET radiomics features in NSCLC is related to overall survival — the "delta radiomics" concept. *Radiother. Oncol.* **118**, S20–S21 (2016).
103. Fave, X. *et al.* Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? *Med. Phys.* **42**, 6784–6797 (2015).
104. Leijenaar, R. T. H. *et al.* The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. *Sci. Rep.* **5**, 11075 (2015).
105. Fave, X. *et al.* Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Sci. Rep.* **7**, 588 (2017).
106. Deasy, J. O. *et al.* Improving normal tissue complication probability models: the need to adopt a "data-pooling" culture. *Int. J. Radi. Oncol. Biol. Phys.* **76**, S151–S154 (2010).
107. Skripcak, T. *et al.* Creating a data exchange strategy for radiotherapy research: Towards federated databases and anonymised public datasets. *Radiother. Oncol.* **113**, 303–309 (2014).
108. Budin-Ljosne, I. *et al.* DataSHIELD: an ethically robust solution to multiple-site individual-level data analysis. *Public Health Genomics* **18**, 87–96 (2015).
109. Schilsky, R. L., Michels, D. L., Kearbey, A. H., Yu, P. P. & Hudis, C. A. Building a rapid learning health care system for oncology: the regulatory framework of CancerLinQ. *J. Clin. Oncol.* **32**, 2373–2379 (2014).
110. MAASTRO clinic. euroCAT: Distributed Learning for Individualized Medicine. *youtube.com*. http://youtu.be/ZDJFOxpwqEA. (2014).
111. The Cancer Imaging Archive. TCIA Collections. *cancerimagingarchive.net* http://www.cancerimagingarchive.net/ (2017).
112. National Cancer Institute, Division of Cancer Treatment & Diagnosis. *Quantitative Imaging Network (QIN)* [online], https://imaging.cancer.gov/programs_resources/specialized_initiatives/qin.htm (2017).
113. Radiological Society of North America. Quantitative Imaging Biomarkers Alliance® (QIBA®). *rsna.org* https://www.rsna.org/qiba/ (2017).
114. QuiC ConCePT. *quic-concept.eu* http://www.quic-concept.eu/ (2017).
115. Benedict, S. H. *et al.* Overview of the American Society for Radiation Oncology–National Institutes of Health–American Association of Physicists in Medicine Workshop 2015: exploring opportunities for radiation oncology in the era of big data. *Int. J. Radi. Oncol. Biol. Phys.* **95**, 873–879 (2016).
116. Meldolesi, E. *et al.* An umbrella protocol for standardized data collection (SDC) in rectal cancer: a prospective uniform naming and procedure convention to support personalized medicine. *Radiother. Oncol.* **112**, 59–62 (2014).
117. EuroCAT Umbrella Protocol for NSCLC. *CancerData.org* https://www.cancerdata.org/resource/doi%3A10.17195/candat.2013.08.1 (2017 May 18th).
118. van Rossum, P. S. *et al.* The incremental value of subjective and quantitative assessment of ^{18}F-FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. *J. Nucl. Med.* **57**, 691–700 (2016).
119. Huang, Y. Q. *et al.* Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J. Clin. Oncol.* **34**, 2157–2164 (2016).
120. Coroller, T. P. *et al.* CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother. Oncol.* **114**, 345–350 (2015).
121. Huynh, E. *et al.* CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. *Radiother. Oncol.* **120**, 258–266 (2016).
122. Cunliffe, A. *et al.* Lung texture in serial thoracic computed tomography scans: correlation of radiomics-based features with radiation therapy dose and radiation pneumonitis development. *Int. J. Radiat. Oncol. Biol. Phys.* **91**, 1048–1056 (2015).
123. Liang, C. *et al.* The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I-II and stage III-IV colorectal cancer. *Oncotarget* **7**, 31401–31412 (2016).
124. Hawkins, S. *et al.* Predicting malignant nodules from screening CT scans. *J. Thorac. Oncol.* **11**, 2120–2128 (2016).
125. Grossmann, P. *et al.* Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in glioblastoma. *BMC Cancer* **16**, 611 (2016).
126. Huang, Y. *et al.* Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology* **281**, 947–957 (2016).
127. Leijenaar, R. T. *et al.* External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. *Acta Oncol.* **54**, 1423–1429 (2015).
128. Cui, Y. *et al.* Quantitative analysis of ^{18}F-fluorodeoxyglucose positron emission tomography identifies novel prognostic imaging biomarkers in locally advanced pancreatic cancer patients treated with stereotactic body radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **96**, 102–109 (2016).
129. Li, H. *et al.* MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. *Radiology* **281**, 382–391 (2016).

Acknowledgments

The authors acknowledge financial support from ERC advanced grant (ERC-ADG-2015, no. 694812) and the QuIC-ConCePT project, which is partly funded by EFPI A companies and the Innovative Medicine Initiative Joint Undertaking (IMI JU) under Grant Agreement no. 115151. This research is also supported by the Dutch Technology Foundation STW (grant no. 10696 duCAT & P14-19 Radiomics STRaTegy), which is the applied science division of NWO, and the Technology Programme of the Ministry of Economic Affairs. Authors also acknowledge financial support from the National Institute of Health (NIH-USA U01 CA 143062–01, Radiomics of NSCLC), EU 7th framework program (EURECA, ARTFORCE – no. 257144, REQUITE – no. 601826), SME phase 2 (EU proposal 673780 – RAIL), the European Program H2020 (BD2Decide – PHC30-689715, ImmunoSABR – no. 733008, PREDICT - ITN no. 766276), Kankeronderzoekfonds Limburg from the Health Foundation Limburg and the Dutch Cancer Society (KWF UM 2011–5020, KWF UM 2009–4454, KWF MAC 2013–6425, KWF MAC 2013–6089) and Alpe d'HuZes-KWF (DESIGN), Center for Translational Molecular Medicine (TraIT), EUROSTARS (SeDI, CloudAtlas, and DART), Interreg V-A Euregio Meuse-Rhine ("Euradiomics") and Varian Medical Systems (VATE and ROO).

Competing interests statement

A.D., leader of the Knowledge Engineering division at MAASTRO, A.J., T.L., J. v- S. and S.W. declare they receive financial support from Varian Medical Systems, a company developing a rapid learning health-care system. R.L. is a salaried employee of, and T.D. consults for ptTheragnostic B.V., a company developing biomarkers and software to individualize radiotherapy treatment. R.T.H.L. and P.L. are co-inventors of radiomics patents (EP2793164A1, US20160203599A1, and WO 2016060557A1).

Publisher's note

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

SUPPLEMENTARY INFORMATION

See online article: S1

ALL LINKS ARE ACTIVE IN THE ONLINE PDF