

Radiomics and Radiogenomics

Karlo Blažetić, mag.bioinf.¹, **Maja Karić**, univ.mag.admin.sanit.^{1,2}

¹Klinički bolnički centar Rijeka, Klinički zavod za radiologiju

²Sveučilište u Rijeci, Fakultet zdravstvenih studija, Katedra za radiološku dijagnostiku

Summary

In the recent past, medicine has not had much choice in treating patients or predicting the course of the disease. Today, modern methods are being developed that will improve the daily life of patients in the form of better diagnostics and therapy.

The idea that is based on personalized medicine by collecting data from medical images is called radiomics.

Radiomics collects data from radiological images such as computed tomography or magnetic resonance imaging. Features collected from the object of interest are shape, texture, intensity, histogram, margins, etc.

These data are subsequently combined with demographic and clinical data to make an analysis. So, radiomics through artificial intelligence, segmentation and other methods may in the future provide the best possible medical care for everyone. The science that links radiological and genetic data is called radiogenomics.

Radiogenomics tends to be able to replace some of the pathohistological methods based on imaging data.

Thus, the main goals of radiomics and radiogenomics are: Prediction and prognosis of the course of the disease, determining the best therapy and providing less invasive diagnostics.

Keywords: Radiomics, Radiogenomics, Data analysis, Deep learning, Personalized medicine, Precise medicine

Introduction

In the beginning of medicine, to be more precise during the ancient civilization, physicians could predict an outcome of disease in oncological patients only as the fatal outcome or short life expectancy. With the discovery of Wilhelm Conrad Röntgen, a revolution in medicine had begun. X-rays have been discovered, which significantly improved the diagnostics of malignant disease, but the therapy was still quite limited. In later years, with the development of chemotherapy, great progress was made in the treatment of oncology patients. In modern times, many other techniques for the treatment and diagnosis of malignant diseases have been developed. There is a problem when we do not know which of these treatments is the best. The answer would be that the best therapy procedure is the one that is the most efficient or the therapy that reduces the tumor the most. Of course, there can be a misinterpretation of the data between physicians. For example, two radiologists can interpret the same CT scans each in their own way.

We need precision medicine to get the right treatment for the patient. One of the subfields of precision medicine is radiomics through which we can train computer system to help us diagnose and choose the best therapy for patients. Also, the computer system would predict the outcome of the disease and prognosis for patients. That

system works on principles of correlation between old and new patient data. In that way, we can decide which therapy is the best for an individual patient.

Precision medicine is a modern field of medicine that optimizes genotype and phenotype characteristics of the individual and its diseases. Furthermore, precise medicine includes system biology, which integrates mathematical modelling and biology, genomics, transcriptomics, proteomics and metabolomics.

Moreover, medical imaging is also ready to be in the center of the action. Radiomics extract quantitative data from many medical images and then it combines it with clinical and patient data in the same database which is also used for the field of radiogenomics. Radiogenomics is a combination of genetic and radiomic data. Radiogenomics can offer voxel by voxel genetic information for a whole heterogenic tumor or in case of metastatic disease, a set of tumors, and then lead to the right individual therapy. Radiogenomics can also quantify characterizations of lesions for better differentiation between benign and malignant entities which helps to carry out better screening of patients [1].

In practice, the job of a radiologist is focused on image interpretation which is based on studying through many years of faculty education. In nuclear medicine, physicians analyze metabolic activity and quantify specific values in the detection of radiation and characterization

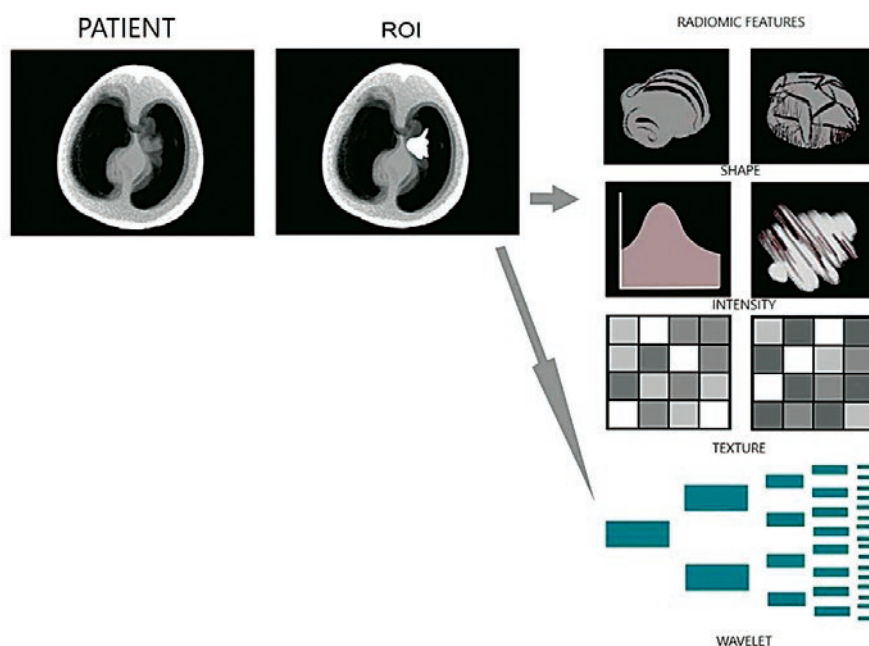


Figure 1: Radiomics analysis pipeline on medical images. CT image with a label „patient“ show a brain cell tumor that is later segmented from the surrounding healthy tissue by ROI. Features are extracted to quantify the tumor shape, intensity, texture and wavelet.

of the object, such as location, shape, sharpness, and intensity. Today, such implications are crucial for diagnostic and therapy procedures, but also for monitoring of the patient, prognosis, etc.

With the correct specification of the occurrence of liver lesions in the application of CT technique with a contrast agent, scans can be used to conclude their molecular properties. Similar to that, some studies indicate that quantitative analysis of grey scale values from images of magnetic resonance imaging (MRI) and scans from patients with oligodendroglioma can be used for the prediction of their genetic marks with deletion of chromosome 11 and 19q [2].

These, as well as earlier examples, show that radiological records can be quantified and shed a whole new light on the diagnosis and development of molecular diagnostics through radiological techniques.

The field of scientific research which connects imaging data and basic genomic data is called radiogenomics or imaging genomics.

Two-way correlations and predictions between imaging features and genomic data can be performed and have been shown to non-invasively identify genomic features and investigate potential biological or molecularly targeted events by analyzing widely available imaging data. There is hope that with the discovery of these characteristics, the diagnostic, predictive and prognostic markers will be used for the decision of optimal steps in the suppression of tumors at an earlier stage, leading to a possibly safer alternative therapy that avoids side effects due to suboptimal treatment.

Radiogenomic research became more feasible in recent years with the appearance of multi-institutional imaging and genomic database records of patients with

cancer such as Atlas Genoma (TCGA), Repository data of molecular brain neoplasia (Rembrandt) and Archive of cancer imaging (TCIA). Research in this field is promising, with all sorts of radiogram features that are connected with the wide spectra of molecular data such as mutations, expression and epigenetic regulation, etc.

Radiomics is defined as the automatic extraction of high, multidimensional features of a quantitative image from standard medical images or in other words radiograms.

Radiomics can extract data from microscopic structures that are not otherwise visualized by the human eye. It can be assumed that such quantitative micro dimensional imaging methods of characterization can more accurately investigate the complex heterogeneity inherent in solid tumors. Considering that one image series could contain millions of voxels which lead to numerous features of size, texture, shape, and wavelet in every subregion, radiomic offers a large amount of imaging features that can be used to facilitate the diagnosis of cancer, staging of the disease, predicting treatment responses and ultimately more precise personalized treatment as shown in Figure 1.

Advances in high-performance computing, computing power, and computational methods have helped ease the foundations for the success of radiomics and quantitative radiogenomics. New advances in software for image segmentation and tools which can automatically, semi-automatically, or manually do image segmentation had provided a mechanism needed for modelling or selection of tumor or tissue of interest which will then be able to automatically start radiomic analysis. To date, no fully automated segmentation tool has been developed to provide sufficient accuracy assessment of tumor shape or tissue

boundaries, and therefore semi-automated and manual segmentations remain the best pathway for accurate segmentation [3]. With segmentation, tumors and its sub-regions can be quantified based on the volume, however there are still different limitations such as the impossibility to get subtle differences and heterogeneity inside the tumor which are not seen by the human eye. Therefore, radiomics is a new field that can help to fill these gaps through the extraction of data and quantitative analysis of macroscopic images.

Basic principles of radiogenomics and radiomics

The working principle of radiogenomics and radiomics is presented in Figure 2.

Conventional radiomics flow combines semantic and human engineering computational features. Those radiomic features can be combined with clinical and demographic data before the last phase of classification which generates results such as benign/malignant, reactive/non-reactive, probability of survival etc. [4].

Radiomics for any set of data can be in 2D, 3D or higher dimension.

Basic components are: Volume of interest (VOI), marking the tissue with semantic properties, segmentation of VOI and calculation of image feature made by human.

Also, collections of imaging features can be combined in vector properties. Now we will describe steps that are needed for radiomic analysis:

VOI identification: Every VOI which is processed needs to be identified first, semi-automatically or manually by a radiologist or automatically with computed aided detection. In imaging of cancer, e.g., when multiple tumors are present in a single imaging study, human effort is generally required to identify those that are clinically significant, as in the case of index lesions with RECIST scores ("criteria for evaluating responses in solid tumors").

Semantic features: Semantic features are descriptive observations of image content. For example, semantic features of a lung tumor may include "the right upper lobe", "pleural attachment", "ground-glass opacity" and so on. There is a problem when radiologists use different terms for the same disease interpretation while evaluating results so the semantic features may not be optimal.

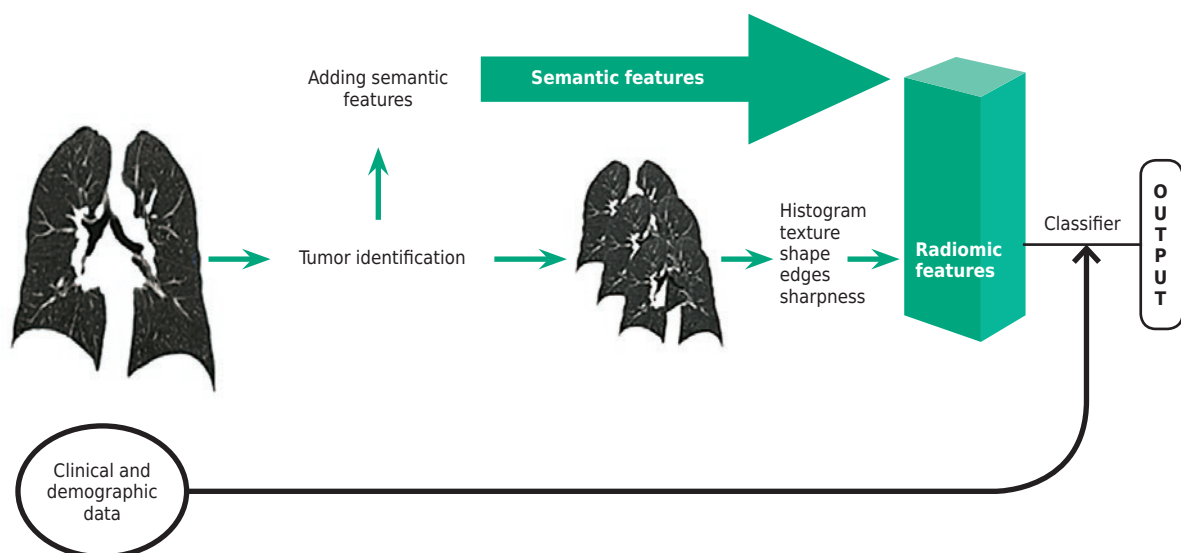


Figure 2: Conventional radiomics workflow

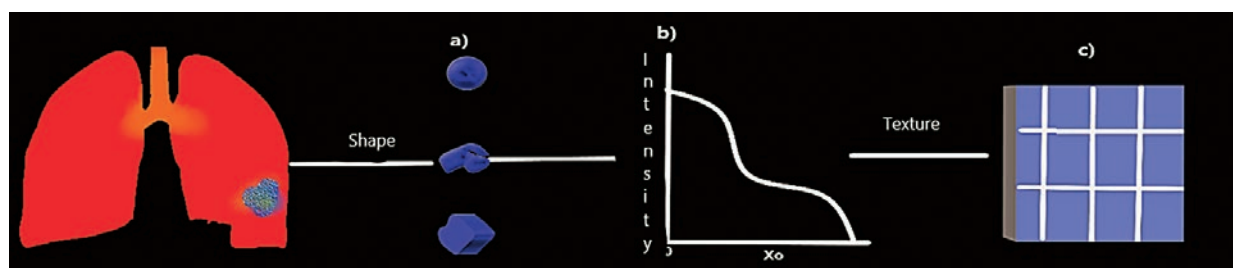


Figure 3: Radiomic features. Schematic representation of lung tumors and determination of features such as shape, intensity and texture for further analysis.

VOI segmentation: Radiomic features can be extracted from inside the volume of the image. In cancer imaging, some parts can contain the whole tumor, subregion and peritumoral region. These features must be identified and entered into algorithms for calculating radiomic features. This step of segmentation is the most difficult of all. Image can be sensitive to the surrounding structures, noise, and lack of information. Because of those technical problems, there is a practice today that every segmentation needs to be checked by a human.

Image feature computation: We can divide features from computed images into several classes: those that describe shape, sharpness, margins, histograms, texture, space variation, etc. which can be seen in Figure 3. Furthermore, inside each, there are many individual features. Many features are mutually connected and as a result, not every feature can provide an independent predictive model. The standard approach in many studies is to generate autocorrelation matrices and combine correlated features in a single descriptor [5].

Methods in radiogenomics

Huge amount of biomedical data characterizing complex diseases are now publicly available, and new computational tools for their quantitative analysis are being routinely developed. None other than radiogenomic field is a better example of the potential in quantitative analysis of biomedical data.

Different measurements such as molecules, tissues, cells, and others are connected so they can be mathematically modelled. For instance, we can see the best example in oncology where we can collect data by routine procedures, and then there is the possibility to connect molecular and imaging data. We can say that radiogenomics binds molecular data from the genomic database and quantitative features of radiological images or radiograms.

The most usual data that are being used for analysis are gene expression, DNA and the number of its copies or data about DNA methylation. All „omics“ (Proteomics, pharmacogenomics, etc.) data can also be combined. From the imaging point of view, CT and MRI scans are being mostly used for data observation because those diagnostic modalities have the best space resolution which offers quantitative extraction of features either collected or extracted by deep learning or radiomic techniques.

One of the main goals of radiogenomics is focused on the idea that medical imaging can relate to molecular biology so that we could better understand patient therapy outcomes and which therapy to choose in general. Radiogenomics showed its potential through predictions of clinical outcomes such as prognosis and prediction of molecular features of the tumor. Some examples of tumor features are receptors of the epidermal growth factor (EGFR), main drug target, etc. Furthermore, radiogenomic maps give us multivariate analysis between image and gene expression or association among image phenotype and potential therapy. Also, CT scans of a thorax can show EGFR mutations. One of the most complex and crucial questions is how much of the molecular composition of complex diseases is reflected in the 3D occurrence of the

disease recorded by radiological images. It is a question that radiogenomics answers more and more every day.

Example of radiomics and radiogenomics in brain tumors

We will describe an example from one study where different imaging phenotypes such as edema and necrosis were segmented by a semi-automated method under the supervision of a radiologist [6].

In this case of a brain tumor or glioblastoma analysis of MRI, images are already segmented.

Normalization of extracted features has been made from contralateral white matter. After that, multiple techniques of preprocessing had been made such as skull bone elimination by use of functional magnetic resonance, imaging and tool for extraction of brain data (FMRIB) and intensity of all images. In the coming phases, we can separate different features based on intensity, texture, shape, size, position and more from various imaging techniques such as T1WI, T2WI, FLAIR, DWI and PWI. Those extracted features can be set off in two categories, semantic and agnostic. Semantic features radiologists usually use to describe the shape, vascularity, necrosis and other, while in radiomics the focus is on agnostically features and mathematically extracted data which are independent of visual criteria.

Agnostic features can be grouped into statistical features in different levels. Features of the first level can be histogram based on one voxel, median, variance, kurtosis, etc. They are derived from regional intensity distribution while second level features include metrics such as entropy, homogeneity, etc. Ultimately, features of higher levels can be collected by evaluation of reciprocal connection of three or more voxels methods such as a performance length matrix or an intensity magnitude zone matrix. The last step in radiomic assessment is biostatistical/bioinformatics analysis.

Although this step includes the full spectrum, from simple descriptive statistics to sophisticated multivariate analysis and machine learning, the multitude of imaging and radiogenomic data is best analyzed by big data machine learning methods and predictive modeling techniques. The feature selection step is an important step in big data analysis, including radiomics and radiogenomics, and is performed prior to predictive modeling/machine learning classification methods.

Glioblastoma is a heterogenic lesion that contains a dynamic cell population with various phenotypes and proliferation characteristics in the same patient, as well as in different patients. Glioblastoma is especially convenient solid tumor for analysis of tumor heterogeneity by use of radiomics. Radiomics which uses advance or more traditional method of MRI has been used for more detailed characterization of certain image phenotypes such as edema and tumor invasion or spatial heterogeneity which are then connected to genomic information with the goal to better understand some of the fundamental mechanisms. A recent study based on the biopsy of image-guided samples collected from the tumor nucleus and surrounding tumor regions measured the ability of MRI texture features to predict spatial histological heterogeneity [7].

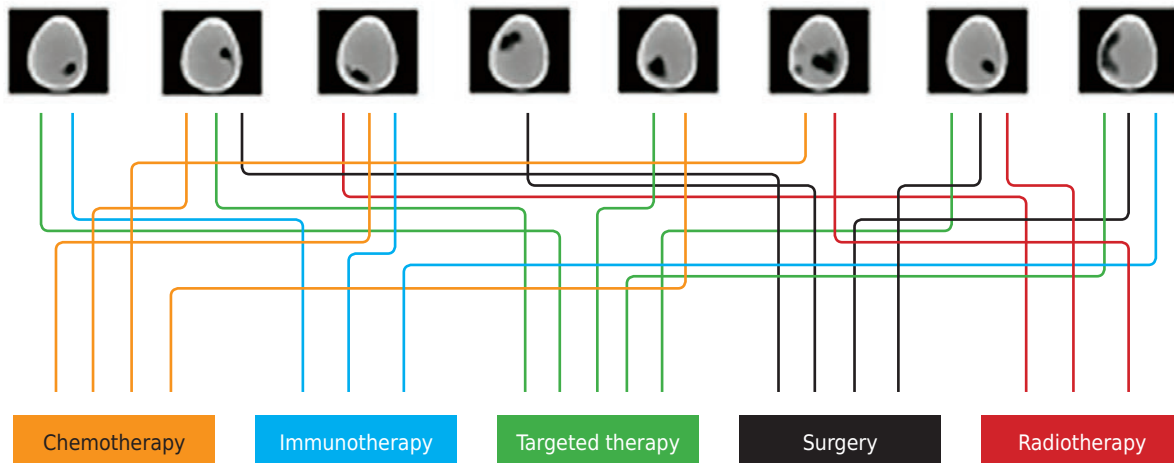


Figure 4: Personalized medicine. Personalized medicine can determine the best possible therapy for an individual. Some of these therapies are immunotherapy, radiotherapy, targeted therapy, surgery and chemotherapy.

In the future, those techniques can be of huge help in high tumor heterogeneity identification which can be very useful for neurosurgeons to eliminate highly infiltrative tumor regions. Another potential application of radiomics relates to post-treatment tumor monitoring.

Currently, histopathological analysis is the golden standard for the diagnosis of brain tumor. However, new studies indicate that noninvasive radiomics correlate with tumor staging and proper diagnosis [8]. Finally, radiomic features relate to genomic information for further development of radiogenomics. In a recent study, radiomic features derived from conventional MRI sequences were able to distinguish TP53, PTEN, and EGFR mutated genes from wild-type glioblastomas [9].

Radiomic analysis should determine the course of the treatment. Thus, radiomics can relate the phenotype obtained from radiological images based on which personalized therapy is determined as shown in Figure 4.

Conclusion

Radiomics and radiogenomics is a new branch of medicine that is still in development. It is based on the collection of various data and the correlation of known data with those that have yet to be analyzed. It is used for predictive models to assess the course and duration of the disease, personalized treatment choices, and as assisted diagnostics. This field of medicine will help medical doctors in their daily work and patients to get the best possible therapy and better outcome in recovery. An individualized approach will significantly shorten the time of diagnosis and implementation of therapy.

References

1. Katja Pinker, MD, PhD2 Joanne Chin, BS, MFA Amy N. Melsaether, MD Elizabeth A. Morris, MD Linda Moy, MD. Precision Medicine and Radiogenomics in Breast Cancer: New Approaches toward Diagnosis and Treatment Radiology. 2018,Jun;287(3): 1-11
2. Brown R, Zlatescu M, Sijben A, Roldan G, Easaw J, Forsyth P, Parney I et al., The use of magnetic resonance imaging to noninvasively detect genetic signatures in oligodendroglioma. Clin Cancer Res, 2008; (14):2357-2362.
3. Fyllingen EH, Stensjøen AL, Berntsen EM, Solheim O, Reinertsen I. Glioblastoma segmentation: Comparison of three different software packages. PLoS One 2016; 11: e0164891.
4. Ruijiang Li, Lei Xing, Sandy Napel, Daniel L. Rubin (2019). Daniel (Daniel L.). Radiomics and radiogenomics : technical basis and clinical applications: 4-8, 203-229.
5. Balagurunathan Y, Kumar V, Gu Y, Kim J, Wang H, Liu Y, Goldgof DB et al. Test-retest reproducibility analysis of lung CT image features. J Digit Imaging, 2014; 27:805-823.
6. Zinn PO, Singh SK, Kotrotsou A, Hassan I, Thomas G, Luedi MM et al. A co-clinical radiogenomic validation study—Conserved magnetic resonance radiomic appearance of periostin expressing glioblastoma patients and xenograft models. Clin Cancer Res 2018. doi:10.1158/1078-0432.CCR-17-3420.
7. PO, Singh SK, Kotrotsou A, Hassan I, Thomas G, Luedi MM et al. A co-clinical radiogenomic validation study—Conserved magnetic resonance radiomic appearance of periostin expressing glioblastoma patients and xenograft models. Clin Cancer Res, 2018. doi:10.1158/1078-0432.CCR-17-3420.
8. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A et al. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol 2007; (114): 97-109.
9. Zinn PO, Singh SK, Kotrotsou A, Abrol S, Thomas G, Mosley J et al. Distinct radiomic phenotypes define glioblastoma TP53-PTEN-EGFR mutational landscape. Neurosurgery 2017; (64): 203-210.