

Research Perspective

Radiomic-based diagnostics in oncology: challenges toward clinical practice

Emanuele Barabino¹, Giovanni Rossi^{2,3}, Alessandro Fedeli⁴, Giuseppe Cittadini⁵ and Carlo Genova^{6,7}

¹Interventional Angiography, Ospedale Santa Corona - ASL 2 Savonese, Pietra Ligure, Italy

²Medical Oncology Department, Ospedale Padre Antero Micone, Genova, Italy

³Department of Medical, Surgical and Experimental Sciences, University of Sassari, Sassari, Italy

⁴Department of Electrical, Electronic, Telecommunications Engineering, and Naval Architecture, University of Genoa, Genova, Italy

⁵UO Radiologia Generale, IRCCS Ospedale Policlinico San Martino, Genova, Italy

⁶UOC Clinica di Oncologia Medica, IRCCS Ospedale Policlinico San Martino, Genova, Italy

⁷Dipartimento di Medicina Interna e Specialità Mediche (DiMI), Facoltà di Medicina e Chirurgia, Università degli Studi di Genova, Genova, Italy

Correspondence to: Carlo Genova, **email:** carlo.genova@hsanmartino.it

Keywords: radiomics; lung cancer; oncology; artificial intelligence; machine learning

Received: April 15, 2021

Accepted: May 24, 2021

Published: June 02, 2021

Copyright: © 2021 Barabino et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Since radiologic exams were converted from analog to digital images, radiologists have made efforts to extract quantitative data from them. In 2012, the word “radiomics” was introduced by Lambin [1] to describe this new science that aims at extracting quantitative features from diagnostic images [2]. One of the possible applications of radiomics in oncology is the investigation of connections between specific radiological phenotypes and molecular status, which in turn translates into personalized therapeutic decisions. Advantages of radiomics are non-invasiveness and the ability to describe pathological processes in their entirety (i.e., the gross tumor volume) and *in-vivo* (e.g., peri-tumoral environment, anatomical relationships), unlike biopsy. The dimensionality of features, the so-called “large p, small n” problem, which is the disproportion between the number of features (p) and images/patients (n) involved in the study, made radiomic features hardly processable with univariate tests, because of the random chance of statistically significant findings [3], or multivariable statistics, due to high correlation between features [4]. Since the amount of information collected through radiomic features makes traditional statistical analyses impractical, they were mostly abandoned in favor of other solutions, such as Artificial Intelligence (AI) and Machine Learning (ML). In order to reduce dimensionality of data but preserve their informative content, prominent roles are played by Principal Component Analysis (PCA) [4] or feature selection by Least Absolute Shrinkage and Selection Operator (LASSO). Notably, ML/AI methods allow

creating flexible predictive models, based on few, almost uncorrelated but reliable features. In the first decade of development, radiomic studies were mostly monocentric (or even “mono-scanner”), based on unstandardized features computed by custom in-house software, required manual segmentation, and lacked external validation, resulting in scarcely reproducible models.

In 2021 our group published a study on *EGFR* mutations in NSCLC [5] that started as a monocentric study and consequently evolved during the revision process. We decided to provide a robust initial features selection using a scarcely applied but effective technique, test & re-test. This technique, based on the repetition of a radiological exam to perform cross-reference on data and eliminate radiomic features considered unreliable, in clinical practice clashes with evident ethical issues due to radiation exposure. However, one radiologic exam can provide repeated images: trans-thoracic lung biopsy. Our first predictive ML model achieved 94% accuracy in the internal validation set, but when we tested it on a public dataset available on The Cancer Imaging Archive, accuracy did not exceed 60%. Two problems became evident: our initial model was too dependent on the training cohort and a proper scaling of features was needed to ensure an optimal use of PCA. Once such problems were solved, we realized that diversity within data is not an issue but a value, as an improved variability forces the predictive model to become less dependent on the original cohort and, therefore, more generalizable. To enhance generalization capabilities, we introduced a small dataset

from another hospital of our province in the training set. As a consequence, accuracy decreased in our cohort to 88.1% but significantly improved in external test sets (76.6% and 83.3%, respectively), leading to the final predictive model.

Nowadays, researchers are sharing radiological images to create larger, diversified, public datasets to test newly developed predictive models and many efforts have been made to standardize radiomic features and to create a shared language [6]. Manual segmentation has been gradually substituted by automatic or AI-based algorithms, which made the process faster and more reproducible. Still, many questions remain open and need answers before proceeding further: investigators perform radiomic analysis to answer a specific question (mutational status, prediction of response to therapy) but the connection between radiomic features and biological variables is not straightforward in most cases. Moreover, the correct harmonization of radiomic data between different centers still represents a challenge, especially with small and heterogeneous datasets. Another issue that may prevent the translation of radiomic analyses into clinical practice is related to the lack of prospective randomized trials.

Nonetheless, radiomics currently represents the top of exploration into radiological images: as the atom has been, in physics, the smallest analyzable measure of matter for quite a time, so the pixel/voxel represents for radiomics. Soon, using new generation CT, PET and MR scanners with improved signal and higher spatial resolution, we will be able to further reduce pixel dimensions, explore new aspects of the radiological image and, hopefully, get some insights on its biological counterpart. The human body or a tumor are not isolate entities, but complex biological processes affected by multiple variables; considering only a part of those variables will give a limited view of the process in its entirety [7]. An example is the activity of immune checkpoint inhibitors (anti-PD1/L1), which is not determined exclusively by the level of expression of the target (which is itself heterogeneous within the same tumor), but also by other immunological checkpoints, or by the expression of these receptors on other tissues [8], or even on the disposition of lymphocytes in the tumor (inflamed, immuno-excluded and immune-desert) [9] which can compromise the activity of the drug itself. As the most promising aspect of radiomics may reside in the capability of describing the complexity of biological processes, it is, perhaps, the best tool at our disposal to photograph tumor heterogeneity. If we succeed in translating radiomics into clinical practice, we will obtain a tool that can answer to new, more complex questions leading to a factual personalized medicine.

CONFLICTS OF INTEREST

Carlo Genova received honoraria from Astra Zeneca, Boehringer Ingelheim, Bristol-Myers-Squibb, Merck Sharp & Dohme, Roche; he also received research

funds from Bristol-Myers-Squibb. Giovanni Rossi received honoraria from Novartis and Janssen. The other authors declare no conflicts of interest.

REFERENCES

1. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012; 48:441–46. <https://doi.org/10.1016/j.ejca.2011.11.036>. PMID:22257792
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016; 278:563–77. <https://doi.org/10.1148/radiol.2015151169>. PMID:26579733
3. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, et al. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012; 30:1234–48. <https://doi.org/10.1016/j.mri.2012.06.010>. PMID:22898692
4. Matzner-Lober E, Suehs CM, Dohan A, Molinari N. Thoughts on entering correlated imaging variables into a multivariable model: Application to radiomics and texture analysis. *Diagn Interv Imaging*. 2018; 99:269–70. <https://doi.org/10.1016/j.diii.2018.04.011>. PMID:29751945
5. Rossi G, Barabino E, Fedeli A, Ficarra G, Coco S, Russo A, Adamo V, Buemi F, Zullo L, Dono M, De Luca G, Longo L, Dal Bello MG, et al. Radiomic Detection of EGFR Mutations in NSCLC. *Cancer Res*. 2021; 81:724–731. <https://doi.org/10.1158/0008-5472.CAN-20-0999>. PMID:33148663
6. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJ, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020; 295:328–38. <https://doi.org/10.1148/radiol.2020191145>. PMID:32154773
7. Saito Y, Horiuchi S, Morooka H, Ibi T, Takahashi N, Ikeya T, Shimizu Y, Hoshi E. Inter-tumor heterogeneity of PD-L1 expression in non-small cell lung cancer. *J Thorac Dis*. 2019; 11:4982–91. <https://doi.org/10.21037/jtd.2019.12.24>. PMID:32030214
8. Daassi D, Mahoney KM, Freeman GJ. The importance of exosomal PDL1 in tumour immune evasion. *Nat Rev Immunol*. 2020; 20:209–15. <https://doi.org/10.1038/s41577-019-0264-y>. PMID:31965064
9. Xiao Y, Ma D, Zhao S, Suo C, Shi J, Xue MZ, Ruan M, Wang H, Zhao J, Li Q, Wang P, Shi L, Yang WT, et al. Multi-Omics Profiling Reveals Distinct Microenvironment Characterization and Suggests Immune Escape Mechanisms of Triple-Negative Breast Cancer. *Clin Cancer Res*. 2019; 25:5002–14. <https://doi.org/10.1158/1078-0432.CCR-18-3524>. PMID:30837276