Editorial

Radiomics to predict response to immunotherapy, bridging the gap from proof of concept to clinical applicability?

Evidence-driven methods and artificial intelligence are becoming game changers in precision medicine. Such revolution is driven from the increasing availability of preclinical and clinical data, with imaging being predominant, enhancing the standardized longitudinal follow up of the patient and the impressive progress of machine learning during the past two decades. As opposed to the biological hypothesis model that uses data to validate the model, machine learning methods refer to a paradigm shift that targets two different objectives. The first refers to the ability of reproducing human behavior on data without necessarily reproducing human behavior. This is typically the case for domains like diagnostic imaging or pathology where numerous recent studies have demonstrated that artificial intelligence methods can outperform human experts [1,2]. The second, and potentially even more promising direction, is the ability to go beyond human abilities. In this context, machine learning methods seek high order correlations of biomarkers/omics that can create statistically significant correlations between data and clinical outcomes and, if feasible, causality.

The analysis of medical images using computational methods is a predominant research tool to reach the above referenced objectives. [3-6] The central idea is to determine automatically a task-specific set of markers that once computational – often called radiome or radiomics – when combined could explain potential clinical outcomes. This concept has attracted increasing attention in medicine since it could potentially provide treatment selection strategies and patient stratification both for the development of new drugs as well in standard clinical practice. The idea of “companion algorithms”, methods able to asses in a continuous manner the performance of drugs and recommend adjustments, is becoming a tractable objective for the years to come. Such a need is even further enhanced by the continuous development of precision medicine drugs targeting smaller and smaller subset of populations. Surgery, radiation therapy [7,8,9] and immunotherapy [11, 12] are domains on which, in recent years, an important number of proof-of-concept studies have been reported demonstrating the potentials of such powerful combination of omics data with advanced mathematical modeling through machine learning.

In the paper by Trebeschi, et al in this issue of *Annals*[13], an approach to predict treatment response for immunotherapy, which is of great promise in oncology, is presented. The materials and methods are heavily in line, both in terms of methodology and clinical setting, with the previous work published by Sun et al [12] where a combined clinical/imaging signature was introduced associated with a computational algorithm to assess treatment response to immunotherapy. The design of the study presented here is quite similar with

the one reported by Sun et al [12], referring to a multi-localization/multi-type tumor cohort and similar imaging biomarkers. It is also well aligned with an important number of recent publications and proof of concept studies demonstrating the interest of investigating imaging biomarkers in similar clinical contexts [14, 15, 16]. The method relies on a predefined set of handcrafted biomarkers in which a three-class classification algorithm (stable, progression, remission) using a conventional machine learning (random forests) is built from a training set after an exhaustive testing of various classifications algorithms. The results are also consistent with those reported by Sun et al and therefore one can conclude that imaging biomarkers and computational algorithms could lead to better patient stratification and provide promising prognostic and predictive tools. On the negative side, performance varies significantly for different tumor localization and tumor types that can be explained using three different hypotheses: (i) insufficient training data, (ii) poor generalization, and (ii) discovering correlations rather than causality. These elements are not a specific limitation to this study but are relevant to almost all recent literature in the field.

The volume of training and testing data is clearly a bottleneck for the development of computational methods in clinical practice. Machine learning algorithms end up solving ill-posed mathematical problems (the number of constraints to solve the problem is fairly low compared to the number of degrees of freedom/parameters to be estimated) and end up being overly sensitive, which heavily compromises their potential impact. The problem could be further amplified when targeting multi-class classification in the presence of tumors with different phenotypes. Generalizability is another limiting aspect on the use of radiomics for treatment response assessment in the context of immunotherapy and beyond that also relates to the quantity and the quality of data as well to the specific characteristics of the machine learning algorithm. Methods that are model-free often lack biological evidence or biological modeling and purely explore evidence-driven approaches that are known to be rather limited in terms of generalization. Last, but not least, establishing causality between imaging biomarkers and biological evidence is mandatory.

In sharp contrast with Sun et al [12], where the model was directly trained on tumor CD8 gene expression and T cell tumor density, the current model is trained and validated using clinical data and images. Secondarily, the authors observe correlation between their signature and genes involved in cell proliferation and mitosis which are not specific to immunological processes. Different machine learning algorithms will end up selecting different features with a rather simple criterion that is the performance on a rather limited validation test. Despite the fact that this paper reports (as well as in [12]) certain biological relevance of the selected features, surprisingly none of these studies agree on the same pool of features and algorithms. Recent difficulties in tackling one single, and rather conventional, biomarker, PDL1 expression, clearly illustrates the huge gap between rational proof of concept and generalization for clinical use [17-19]. Considering radiomics, we will have to integrate much more than just one single parameter into our decision processes.

Radiomics combined with artificial intelligence bears great promises in oncology from a research perspective, and their potential impact can be accelerated through:

- Standardization of acquisition protocols and increasing availability of high quality data. Imaging, and beyond (genomics, phenomics, pathomics), often are acquired using different protocols and clinical settings that may be a bottleneck in performing multi-center translation studies and augmenting the pool of patients that can be used for training these complex algorithms
- Integration of domain knowledge, biological evidence and hypotheses: Evidence-driven approaches bear great potential but lack interpretabilityr robustness and generalization. The ability to combine evidence-driven approaches (bottom up) with domain-driven hypothesis (top-down) will be a game changer in the clinical adoption of these methods
- Integration of multi-modal data at different mid-points (multi-omics/ longitudinal data): Computational imaging bears great promises in oncology as it refers to a standard clinical practice, can be easily acquired and is not invasive. However, it cannot encapsulate on its own the complex behavior of tumor cells, and therefore the combination with other clinical data like genomics, pathomics, etc. – at least during the training of these algorithms – is a necessity.

To conclude, the paper by Trebeschi, et al is of extreme interest, enhancing existing hypotheses on the relevance of radiomics for treatment selection and prognosis in the context of immunotherapy and beyond. At this time, converging data from various groups demonstrate the power of radiomics to predict response to immunotherapy. To move beyond the proof of concept, significant effort must be carried out in translating such research into clinical practice, to guarantee generalization, explicability, and secure fairness of the decision process.

Eric Deutsch[1, 2, 3] & Nikos Paragios[4] *

[1] Department of Radiation Oncology, Gustave Roussy Cancer Campus, 94805 Villejuif CEDEX, France
[2] INSERM 1030 Molecular Radiotherapy, Gustave Roussy Cancer Campus, 94805 Villejuif CEDEX, France
[3] Université Paris-Saclay, Faculté de medecine,94270, le Kremlin-Bicêtre, France
[4] TheraPanacea / Ecole CentraleSupelec / University of Paris Saclay, 75014, Paris CEDEX, France

(*E-mail: nikos.paragios@centralesupelec.fr)

Disclosures

Dr. DEUTSCH reports grants and personal fees from ROCHE GENENTECH, grants from SERVIER, grants from ASTRAZENECA, grants and personal fees from MERCK SERONO, grants from BMS, grants from MSD, outside the submitted work; .

Pr Paragios is CEO and Founder of Therapanacea

References

[1] A. Esteva, B. Kuprel, R. Novoa, J. Ko, S. Swetter, H. Blau & S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. Nature volume 542, pages 115–118 (02) 2017.

[2] H. Haenssle, C, Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo, A, Ben Hadj Hassen, L. Thomas, A. Enk & L. Uhlmann. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Annals of Oncology, Pages 1836–1842, Volume 29, Issue 8, August 2018.

[3] N. Paragios, J. Duncan & N. Ayache. Handbook of Biomedical Imaging: Methodologies and Clinical Research. Springer, ISBN 978-0-387-09749-7, 2015.

[4] H. Aerts, E. Rios Velazquez, R. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M. Rietbergen, C. René Leemans, A. Dekker, J. Quackenbush, R. Gillies & P. Lambin. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications volume 5, Article number: 4006 (2014).

[5] R. Gillies, P. Kinahan & H. Hricak. Radiomics: Images Are More than Pictures, They Are Data. Radiology. Pages: 563-77, Feb;278(2), 2016.

[6] E. Limkin, R. Sun, L. Dercle, E. Zacharaki, C. Robert, S. Reuzé, A. Schernberg, N. Paragios, E. Deutsch & C. Ferté. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Annals of Oncology, Pages 1191–1206, Volume 28, Issue 6, June 2017.

[7] J Lötsch , R Sipilä,T Tasmuth, D Kringel, AM Estlander, T Meretoja, E Kalso, A Ultsch . Machine-learning-derived classifier predicts absence of persistent pain after breast cancer surgery with high accuracy. Breast Cancer Res Treat. Pages 399-411, Volume 171, Issue 2 Sept 2018.

[8] M Bahl, R Barzilay, AB Yedidia, NJ Locascio, L Yu, CD Lehman . High-Risk Breast Lesions: A Machine Learning Model to Predict Pathologic Upgrade and Reduce Unnecessary Surgical Excision. Radiology. Pages 810-818, Volume 286, Issue 3, Mar 2018.

(9) T. Coroller, P. Grossmann, Y. Hou, E. Rios Velazquez, R. Leijenaar, G. Hermann, P. Lambin, B. Haibe-Kains, R. Maka & H. Aerts. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiotherapy and Oncology. Pages 345-350, Volume 114, Issue 3, March 2015.

(10) S. Reuzé, A. Schernberg, F. Orlhac R. Sun, C. Chargari, L. Dercle, E. Deutsch, I. Buvat & C. Robert. Radiomics in Nuclear Medicine Applied to Radiation Therapy: Methods, Pitfalls, and

Challenges. International Journal of Radiation Oncology, Pages 1117-1142, Volume 102, Issue 4, 15 November 2018.

[11] S Chen, S Feng, J Wei, F Liu,B Li, X Li, Y Hou, D Gu, M Tang, H Xiao, Y Jia, S Peng, J Tian, M Kuang. Pretreatment prediction of immunoscore in hepatocellular cancer: a radiomics-based clinical model based on Gd-EOB-DTPA-enhanced MRI imaging. Eur Radiol. 2019 Jan 21.

[12] R. Sun, E. Limkin, M. Vakalopoulou, L. Dercle, S. Champiat, S. Rong Han, L. Verlingue, D. Brandao, A. Lancia, S. Ammari, A. Hollebecque, J-Y. Scoazec, A. Marabelle, C. Massard, J-C. Soria, C. Robert, N. Paragios, E. Deutsch, C. Ferté. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. Lancet Oncology, pages 1180-1191, Volume 19, ISSUE 9, September 2018.

[13] S. Trebeschi, S.G. Drago, N.J. Birkbak, et al. Predicting response to cancer immunotherapy using non-invasive radiomic biomarkers. Ann Oncol 2019; 30: doi.org/10.1093/annonc/mdz108

[14] B. Larimer, E. Wehrenberg-Klee, F. Dubois, A. Mehta, T. Kalomeris, K. Flaherty, G. Boland & U. Mahmood. Granzyme B-PET Imaging as a Predictive Biomarker of Immunotherapy Response. Cancer Research, 1;77(9):2318-2327, May 2017.

(15)C Tang, B Hobbs, A Amer, X Li, C Behrens, JR Canales, EP Cuentas, P Villalobos, D Fried, JY Chang, DS Hong, JW Welsh, B Sepesi, L Court, II Wistuba, EJ Koay. Development of an Immune-Pathology Informed Radiomics Model for Non-Small Cell Lung Cancer. Sci Rep. Page 1992,Volume 8, Issue 1, Jan 2018.

(16) I Tunali, JE Gray, J Qi, M Abdalah, DK Jeong, A Guvenis, RJ Gillies , MB Schabath. Novel clinical and radiomic predictors of rapid disease progression phenotypes among lung cancer patients treated with immunotherapy: An early report. Lung Cancer. Pages 75-79, Volume 129, Mar 2019.

(17) S. Hendry, D. Byrne, G. Wright, R. Young, S. Sturrock, W. Cooper & S. Fox. Comparison of Four PD-L1 Immunohistochemical Assays in Lung Cancer. Journal of Thoracic Oncology. Pages 13(3), March 2018.

[18] J. Sunshine, P. Nguyen, G. Kaunitz, T. Cottrell, S. Berry, J. Esandrio, H. Xu, A. Ogurtsova, K. Bleich, T. Cornish, E. Lipson, R. Anders& J-M. Taube JM. PD-L1 Expression in Melanoma: A Quantitative Immunohistochemical Antibody Comparison. Clinical Cancer Research. Pages 4938-4944, 15;23(16), August 2017.

[19] D. Rimm, G. Han, J-M Taube, E. Yi, J. Bridge, D. Flieder, R. Homer, W West, H. Wu, A. Roden, J. Fujimoto, H. Yu, R. Anders, A. Kowalewski, C. Rivard, J. Rehman, C. Batenchuk, V. Burns, F. Hirsch, I. Wistuba. A Prospective, Multi-institutional, Pathologist-Based Assessment

of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. JAMA Oncology, pages 1051-1058, 1;3(8), August 2017.