

Author response to Cunha *et al*

Rivka R Colen,^{1,2} Christian Rolfo,³ Murat Ak,^{1,2} Mira Ayoub,^{1,2} Sara Ahmed,⁴ Nabil Elshafeey,⁵ Priyadarshini Mamindla,² Pascal O Zinn,⁶ Chaan Ng,⁷ Raghu Vikram,⁷ Spyridon Bakas,⁸ Christine B Peterson,⁹ Jordi Rodon Ahnert,¹⁰ Vivek Subbiah,¹⁰ Daniel D Karp,¹⁰ Bettzy Stephen,¹⁰ Joud Hajjar,^{11,12} Aung Naing ¹⁰

To cite: Colen RR, Rolfo C, Ak M, *et al*. Author response to Cunha *et al*. *Journal for ImmunoTherapy of Cancer* 2021;**9**:e003299. doi:10.1136/jitc-2021-003299

Accepted 08 July 2021

ABSTRACT

The need to identify biomarkers to predict immunotherapy response for rare cancers has been long overdue. We aimed to study this in our paper, ‘Radiomics analysis for predicting pembrolizumab response in patients with advanced rare cancers’. In this response to the Letter to the Editor by Cunha *et al*, we explain and discuss the reasons behind choosing LASSO (Least Absolute Shrinkage and Selection Operator) and XGBoost (eXtreme Gradient Boosting) with LOOCV (Leave-One-Out Cross-Validation) as the feature selection and classifier method, respectively for our radiomics models. Also, we highlight what care was taken to avoid any overfitting on the models. Further, we checked for the multicollinearity of the features. Additionally, we performed 10-fold cross-validation instead of LOOCV to see the predictive performance of our radiomics models.

We appreciate the authors’ interest and commentary to highlight this important paper regarding the ability of radiomics to predict immunotherapy response in advanced cancer. As we know, there are multiple proven methods used in machine learning for radiomics analysis. Various feature selection methods and prediction models can be employed and yield sound data analysis and precise, accurate, and robust results. In our study, we have developed a pipeline for bioinformatic analysis that has been validated both preclinically and clinically for the processing of radiomics data¹ derived from both MRI and CT imaging.

The commenting authors mention, ‘the investigators may adjust the algorithm’s hyperparameters and try again until satisfactory performance is achieved. Since many changes are made to make the model more accurate for the validation data, overfitting may occur.’ While some investigators building tree-based classification algorithms (decision trees, random forests, and eXtreme Gradient Boosting (XGBoost)) use hyperparameter tuning to minimize a certain type of loss function or the classification error rate, we have not. In our case, while building our

radiomics models, we have not performed any grid search for tuning of the algorithm’s hyperparameters. XGBoost parameters are divided into general parameters (booster), booster parameters (eta, min_child_weight, max_depth, max_leaf_nodes, gamma, max_delta_step, subsample, etc), and learning task parameters (objective, eval_metric, and seed).² Keeping in mind the small size of our dataset, we relied on the default values for major parameters from the XGBoost algorithm (booster=gbtree, eta=0.3, max_depth=6, objective = “binary logistic”, and eval_metric=error).

The commenting authors also bring up sample size, tumor type heterogeneity, feature selection stability, and the use of leave-one-out cross-validation (LOOCV). As we acknowledge in our original article, a limitation of our study is the small sample size (N=57); however, we were still able to robustly predict immunotherapy response, thus demonstrating the feasibility of this method in advanced cancer. While we acknowledge the tumor type heterogeneity inherent in advanced rare cancers and while not every tumor histological subtype had responders, the advanced tumor group as per compliance with standard United States Food and Drug Administration (FDA) clinical trial protocol demonstrated that key radiomics features irrespective of tumor type predicted response to immunotherapy. Although it was not the primary aim of this study, we are able to see that tumor as a whole harbor specific radiomics features irrespective of tumor type that can robustly help with patient stratification. Regarding feature selection and predictive modeling, feature selection gains particular importance while working with high dimensional datasets where small-n-large-p problems exist; and choosing the best feature selection method to overcome this problem is crucial. The LASSO feature selection method addresses the small-n-large-p issues by



▶ <http://dx.doi.org/10.1136/jitc-2021-003044>



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Aung Naing;
aanaing@mdanderson.org

Dr Rivka R Colen;
colenrr@upmc.edu

applying a shrinkage (regularization) process,³ through penalizing the coefficients of the regression variables and shrinking some of them to zero.⁴ In doing so, this method reduces the variance without any significant increase of the bias; and hence, this is most useful in cases where there are small number of observations and many (radiomics) features. Moreover, the LASSO feature selection method helps eliminate irrelevant features not associated with the response variable, reduces overfitting, and successfully handles multicollinearity.⁵ Furthermore, LASSO, as a regularization method, is known to handle multicollinearity very well in small datasets when compared with other feature selection methods.⁶ The XGBoost, a tree boosting algorithm used in machine learning, is the most widely recognized classification algorithm and is extensively used as it enables one to do the parallel computation, cross-validation, regularization, tree pruning, and missing value imputation (if needed).² For building our radiomics classification models, we have chosen an ensemble modeling approach where we used least absolute shrinkage and selection operator (LASSO) followed by XGBoost for feature selection and model building, respectively. Feature selection methods like minimum redundancy maximum relevance method and wrapper methods (forward and backward elimination) have been previously tested and implemented as a part of validating our pipeline during its development phase, and we have noted that the LASSO feature selection method resulted in higher predictive accuracy and the selection of more

Table 1 Results for multicollinearity on the top 10 features using the variance inflation factor (VIF). A maximum VIF of 10 is chosen as the threshold and removing features with VIF greater than the threshold is said to reduce the multicollinearity. We see from table 1 that the features P_F269 and P_F289 have VIF >10, meaning these predictors are highly correlated with the remaining features in the model

Top 10 features used for the prediction of RECIST response		Top 10 features used for the prediction of irRECIST response	
Feature	VIF	Feature	VIF
P_F7	1.36	P_F89	1.57
P_F84	1.49	P_F187	1.21
P_F269	15.12	P_F269	12.78
P_F289	15.40	P_F289	12.01
PV_F35	5.38	PV_F24	1.38
PV_F89	5.22	PV_F81	1.27
PV_F109	9.19	PV_F270	1.35
Ar_F152	1.09	Ar_F152	1.35
Ar_F154	1.16	ArV_F36	4.13
Ar_F238	1.17	ArV_F247	3.74

RECIST, Response Evaluation Criteria in Solid Tumors
irRECIST, immune-related Response Evaluation Criteria in Solid Tumors

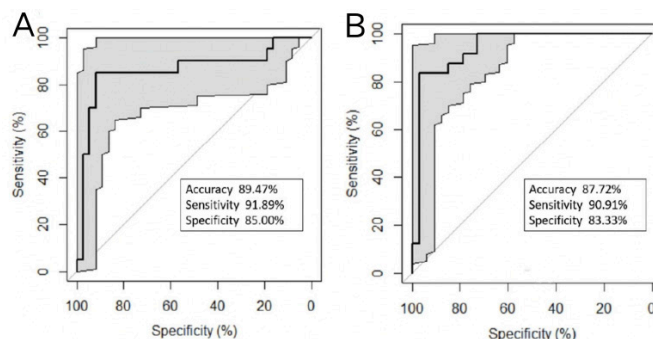


Figure 1 (A) Receiver operating characteristic (ROC) curve representing the performance of the predictive model using top 10 radiomic features to predict RECIST response with 10-fold cross-validation over 10 iterations. (B) ROC curve representing the performance of the predictive model using top 10 radiomic features to predict irRECIST response with 10-fold cross validation over 10 iterations.

meaningful features when compared with others and was best in handling multicollinearity of the features.

Leave-one-out cross-validation is an extension of the k-fold cross-validation, where k is equal to the number of samples in the dataset. LOOCV is best when working with small datasets and when the estimation of the model performance is critical.⁷ Since our dataset is small (N=57 patients), we applied LOOCV as the cross-validation while building our radiomics models. Multiple studies have shown that LASSO for feature selection and XGBoost for classification offer good performance in predictive modeling.⁸ LASSO feature selection, XGBoost, and LOOCV were also used and confirmed in the landmark paper published by Zinn *et al*¹ establishing a causal linkage between radiomics and genomics. Furthermore, this method and pipeline have been validated in both MRI and CT.¹ In addition, immune-related Response Evaluation Criteria in Solid Tumors (irRECIST) adapted the concept of measurement similar to that of Response Evaluation Criteria in Solid Tumors (RECIST). So, we expect to have shared radiomics features extracted in patients assessed by irRECIST and RECIST in our radiomics models.

As per the authors' suggestion, we have checked the multicollinearity of the features and performed 10-fold cross-validation as an alternative to LOOCV to see the predictive performance of our radiomics models. The results for multicollinearity identified by variance inflation factor (VIF) on the top 10 radiomics features used to build the RECIST and irRECIST radiomic models are given in table 1. High VIFs (>10) indicate multicollinearity; this makes interpretation of the contribution of an individual feature challenging but does not necessarily impact predictive performance. Tree-based algorithms such as XGBoost are particularly robust to multi-collinearity. Furthermore, using the 10-fold cross-validation over 10 iterations, the predictive model using the top 10 out of the total 44 features identified from LASSO feature selection to predict RECIST response

resulted in a radiomics model that had an accuracy, sensitivity, and specificity (89.47%, 91.89%, and 85%, respectively) (figure 1A). Similarly, the predictive model using the top 10 out of the total of 56 features identified after LASSO feature selection to predict irRECIST response with 10-fold cross-validation over 10 iterations resulted in a radiomics model that had an accuracy, sensitivity, and specificity (87.72%, 90.91%, and 83.33%, respectively) (figure 1B).

In conclusion, we thank the authors for their comments, and we feel there are several ways to look at the radiomics and our approach has been one of these approaches. More studies are needed to solidify and validate the findings.

Author affiliations

¹Department of Radiology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

²Department of Radiology, Hillman Cancer Center, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA

³Center for Thoracic Oncology at Tisch Cancer Institute, Mount Sinai Health System & Icahn School of Medicine at Mount Sinai, New York, New York, USA

⁴Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

⁵Department of Breast Imaging, Diagnostic Radiology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

⁶Department of Neurosurgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA

⁷Abdominal Imaging Department, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

⁸Department of Radiology, Pathology, and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

⁹Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

¹⁰Investigational Cancer Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

¹¹William T Shearer Center for Human Immunobiology, Texas Children's Hospital, Houston, Texas, USA

¹²Section of Immunology, Allergy and Retrovirology, Baylor College of Medicine, Houston, Texas, USA

Twitter Murat Ak @drmuratak, Nabil Elshafeey @naelshafeey@mdanderson.org and Aung Naing @AnaingMD

Acknowledgements We thank the patients and their families and caregivers for participating in the study.

Contributors RRC, CR and AN contributed to conception and design, provision of study materials or patients, collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. MAK, MAyoub, SA, NE, PM, POZ, CN, RV, SB, CP, JRA, VS, DDK, BS and JH contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work.

Funding This study was funded by National Institutes of Health/National Cancer Institute (grant number: P30CA016672), University of Pittsburgh Hillman Cancer Center (RRC) (grant number: P30CA047904), the University of Texas MD Anderson Cancer Center Institutional Research Grant (IRG), Merck Sharp, and Dohme.

Competing interests CN reports grant support and personal fees from General Electric Healthcare, outside the submitted work; SB reports grant support from National Institutes of Health, outside the submitted work; JRA reports personal fees from Novartis, Eli Lilly, Orion Pharmaceuticals, Servier Pharma, Peptomyc, and

Merck Sharpe, on the advisory board for Novartis, Eli Lilly, Orion Pharmaceuticals, Servier Pharma, Peptomyc, Merck Sharpe and Dome, Kelun Pharma/Klus Pharma, Pfizer, Roche Pharma, and Elipses Pharma, research funding from Bayer, Novartis, Spectrum Pharmaceuticals, Tocagen, Symphogen, BioAtla, Pfizer, GenMab, CytomX, KELUN-BIOTECH, Takeda-Millennium, GLAXOSMITHKLINE, and Ipsen, from null, outside the submitted work; VS reports clinical trial research funding from Novartis, Bayer, GlaxoSmithKline, Nanocarrier, Vegenics, Celgene, Northwest Biotherapeutics, Berghealth, Incyte, Fujifilm, Pharmamar, D3, Pfizer, Multivir, Amgen, Abbvie, Alfa-sigma, Agensys, Boston Biomedical, Idera Pharma, Inhibrx, Exelixis, Blueprint medicines, Loxo oncology, Takeda and Roche/ Genentech, National Comprehensive Cancer Network, NCI-CTEP, and UT MD Anderson Cancer Center, outside the submitted work; JH reports grants from Immune Deficiency Foundation, Jeffrey Modell Foundation and Chao Physician-Scientist, and Baxalta, and has served as an advisory board member for Takeda, CSL Behring, and Horizon Pharma outside the submitted work; AN reports research support and non-financial support from Merck Sharp and Dohme, grants from NCI/NIH, research support from the University of Texas MD Anderson Cancer Center, during the conduct of the study; grants from NCI, research support from EMD Serono, MedImmune, Healios Onc. Nutrition, Attercor, Amplimmune, ARMO BioSciences, Karyopharm Therapeutics, Incyte, Novartis, Regeneron, Merck, Bristol Myers Squibb, Pfizer, CytomX Therapeutics, Neon Therapeutics, Calithera BioSciences, TopAlliance BioSciences, Eli Lilly, Kymab, PsiOxus, Arcus Biosciences, NeolmmuneTech, ImmuneOncia, and Surface Oncology, non-financial support for travel and accommodation from ARMO BioSciences, and has served as an advisory board member for Novartis, CytomX Therapeutics, Genome and Company, STCube Pharmaceuticals, OncoSec KEYNOTE-695, and Kymab, outside the submitted work. RRC, CR, MAK, MAyoub, SA, NE, PM, POZ, RV, CP, BS, and DDK declare no competing interests.

Patient consent for publication Not required.

Provenance and peer review Commissioned; internally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Aung Naing <http://orcid.org/0000-0002-4803-8513>

REFERENCES

- Zinn PO, Singh SK, Kotrotsou A, et al. A Coclinical Radiogenomic validation study: conserved magnetic resonance radiomic appearance of Periostin-Expressing glioblastoma in patients and xenograft models. *Clin Cancer Res* 2018;24:6288–99.
- Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on knowledge discovery and data mining*. San Francisco, California, USA: Association for Computing Machinery, 2016: 785–94.
- Fonti V, Belitser E. *Paper in business analytics feature selection using LASSO*, 2017.
- Tibshirani R. Regression shrinkage and selection via the LASSO: a retrospective. *J R Stat Soc Series B* 2011;73:273–82.
- Gold D. *Dealing with Multicollinearity: a brief overview and introduction to tolerant methods*, 2017.
- Gafar Matanmi Oyeyemi EOO, Folorunsho AI. On performance of shrinkage methods – a Monte Carlo study. *Int J Stat* 2015;5:72–6.
- Cheng H, Garrick DJ, Fernando RL. Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction. *J Anim Sci Biotechnol* 2017;8:38.
- Sun P, Wang D, Mok VC, et al. Comparison of feature selection methods and machine learning classifiers for Radiomics analysis in glioma grading. *IEEE Access* 2019;7:102010–20.