Review of: "Study on the prognosis predictive model of COVID-19 patients based on CT radiomics"

Stefan Pszczolkowski[1]

1 University of Nottingham

Potential competing interests: The author(s) declared that no potential competing interests exist.

This article describe the use of radiomic features extracted from pulmonary CT images to predict outcome (improvement vs aggravation) in patients with COVID-19. Images acquired at one institution were used to extract radiomic features. Then, models using these radiomic features, only clinical features, or the combination of both were compared.
As expected, the models using a combination of features outperformed the models using only radiomics or clinical features, with good overall prediction performance in an external validation set of 64 patients from a different institution. Five different machine learning methods were tested for this purpose and all seem to provide similarly good results.

Despite the overall good results, I have a few comments regarding weaknesses of this work:

- As stated by the authors, the sample size is quite small and it comprises only one training and one testing institution.
- The motivation seems a little weak to me. I can understand that having an accurate prognosis of outcome can better inform treatment, but I am not completely sure that, in practical terms, going through the hassle of acquiring CT images on COVID-19 patients as opposed to, for example, using chest X-rays would be outweighed by the use of a radiomics-based predictive model.
- It is not at all clear how the final set of 19 features were selected. This must be appropriately justified. According to Figure S1, the contributions of the selected features vary greatly, with some of them having regression coefficients very close to zero.
- There is no discussion about the clinical interpretability of the best radiomic features. Why are these 19 features the best ones as opposed to any other set of features? For example, if a feature based on the Laplacian of Gaussian filter is important, it might say something about the heterogeneity of the region, which in turn may have a clinical interpretation.
- The authors mention that resampling the images to 1.0mm isotropic would solve the issue of using different scanners and/or parameterisations. This is not true at all as, for example, the heterogeneity introduced by using different slice thicknesses would just be propagated in the resampling. If the authors plan a larger multicentre study they should consider the use of harmonisation techniques (e.g.

ComBat or domain adaptation) to minimise the biases introduced by different scanners and/or parameterisations.