Journal of Cardiovascular Computed Tomography 12 (2018) 210–211

Contents lists available at ScienceDirect

Journal of Cardiovascular Computed Tomography

journal homepage: www.elsevier.com/locate/jcct

Editorial

Machine learning for predicting death and heart attacks from CCTA

Machine learning (ML) techniques have become commonplace in our daily lives. Prediction of bank customer abilities to pay off loans, driver-assist technologies or shopping suggestions in online shopping ads are currently employing various ML algorithms. In medicine, ML promises a more optimal patient-specific selection of treatment based on the optimal prediction of outcome disease. In cardiac imaging, a few recent reports have applied ML in SPECT and PET myocardial perfusion imaging for diagnostic and prognostic prediction.[1–4] ML with coronary CT angiography (CCTA) data have been proposed for automatic lesion detection,[5] determination of image quality,[6] prediction of death,[7] and prediction of invasive fractional flow reserve (FFR),[8–10] or PET[11] results. Although not yet used clinically, ML techniques are likely to play an essential role in the detection and management of obstructive coronary artery disease (CAD) in the future.

In this month's issue of JCCT, van Rosendael and colleagues apply ML techniques to predict all-cause death or myocardial infarction by training and testing the algorithm in the CONFIRM registry cohort of 8844 patients without known CAD in patients who underwent a CCTA. As input, the algorithm uses 35 CCTA variables visually defined on CCTA during clinical reporting in approximately 12 international centers that are participating in the registry.[12] These features include qualitative 17-segment scoring of stenosis severity and plaque composition. The authors developed the ML score in 80% of the patients, with 20% of the population left aside. They compared the new score to five currently established CCTA scores such as the modified Duke prognostic index, CCTA Leaman score, segment stenosis score (SSS), segment involvement score (SIS) and traditional CAD classification. Unlike previous work with ML in CONFIRM registry (Motwani et al.[7]), they did not use clinical variables for their ML prediction. Their ML score appears to outperform the standard CCTA scores with about a 10% increase in the area under the reciever operator characteristics curve (AUC) in prediction of death and myocardial infarction at any time during the approximately 5-year average follow-up time (the exact minimum and maximum follow up times are not provided).

Possible applications. How could such new ML score be used clinically? An ML model could be provided to clinicians, for example in the form of web application. The developed score is computed from the 35 CCTA variables derived during clinical reporting, based on subjective physician interpretation. The performance of this score, however, would depend on the quality and accuracy of the visual 17-segment CCTA scoring. Meaningful thresholds for interpretation of the result would also need to be established. The risk categories of a new score (for example low risk, high risk) should be matched to those of the already established scores. Alternatively, the continuous probability of the 5-year death or infarct (average follow up time) could be directly presented to the physician. Since the authors did not use clinical variables (such as age and gender) for the derivation of the score, such probability, based on image data alone could be, however, misleading. The likelihood of the 5-year all-cause death will be obviously dramatically different for a 90-year old male as compared to a 50-year old female patient. Thus, if the goal of the new score is the accurate death prediction for clinical purposes, it would be preferable to also incorporate the readily available clinical information in addition to the CCTA variables.

Validation. One of the most critical aspects of ML is the validation regimen. It is of crucial importance that data used for training of the ML models are not used for testing. Indeed, some of the ML methods (for example random forest approaches) allow such flexible models, that they would, in fact, allow obtaining perfect prediction (AUC = 1.0) if the same data were used for training and testing, even in a large number of cases. In general, most non-linear ML techniques are more prone to overfitting than traditional more rigid logistic regression methods. Because of this, if any information is shared between testing and validation data, the estimate of performance is likely to be overly optimistic. In the current study, the authors set aside randomly 20% of the data for the final comparison with the established scores. The 20% consisted of approximately of 1769 datasets and 121 events (assuming stratification of events). While the authors refer to the cross-validation technique, the final score was tested only once on the remaining 20% of the data; therefore, in effect, their evaluation is equivalent to traditional 80/20 split sample. The downside of such split sample techniques is the increased variability of the AUC estimate due to the low number of events in the final testing population, and random selection of the hold-out sample. The cross-validation technique involves repeated testing of multiple models and subsequently concatenation (or averaging) of the results from several models. Such repeated evaluation with multiple models allows for a more precise estimate of the expected performance.

External validation. Another aspect of the validation is the need for the external validation, which authors elude to in the limitations of their work. Indeed, to estimate the “real world” performance of the new ML score in the clinical application, the score should be ideally tested in data from other centers, whose data were not used for the derivation of the score. The external testing would reflect the expected performance of the new score in wider clinical practice-in sites outside of the CONFIRM registry. Although the CONFIRM database is a multicenter registry, one could expect that the estimation of the prediction accuracy of the presented ML approach is more optimistic than that for the compared traditional measures (established without explicit optimization to detect death or myocardial infarction and derived with data from other centers). For the ML score to be widely applicable, the data

https://doi.org/10.1016/j.jcct.2018.05.008

Available online 05 May 2018

in the new unseen centers should be nevertheless similar to the data on which the algorithm was trained. Visual interpretation of the CCTA scan should also be similarly consistent in the new centers. It would not be optimal (or even sensible) to train the ML algorithms on "apples" and apply it to "oranges." Therefore, for the optimal ML deployment, the training population should reflect the variety of patients and results expected in the broader clinical world.

ML has great potential for application in the analysis of the CCTA data. It can aid physicians in providing the most likely diagnosis or prognostic outcome. Future methods should incorporate automated image analysis methods or direct machine learning from images to remove visual subjectivity from the ML prediction. Methods focused on prognostic prediction should include all available clinical data such as gender or age to provide precise patient-specific estimates. Thorough testing, using repeated cross-validation with internal and if possible also with external data, should be used for the evaluation of such techniques.

References

1. Betancur J, Otaki Y, Motwani M, et al. Prognostic value of combined clinical and myocardial perfusion imaging data using machine learning. *JACC Cardiovasc Imaging.* 2017 Oct 16. http://dx.doi.org/10.1016/j.jcmg.2017.07.024 pii: S1936-878X(17) 30804-5. [Epub ahead of print].
2. Haro Alonso D, Wernick MN, Yang Y, Germano G, Berman DS, Slomka P. Prediction of cardiac death after adenosine myocardial perfusion SPECT based on machine learning. *J Nucl Cardiol.* 2018.
3. Betancur J, Commandeur F, Motlagh M, et al. Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study. *JACC Cardiovasc Imaging.* 2018.
4. Arsanjani R, Dey D, Khachatryan T, et al. Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. *J Nucl Cardiol.* 2015;22(5):877–884.
5. Kang D, Dey D, Slomka PJ, et al. Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography. *J Med Imaging.* 2015;2(1) 014003.
6. Nakanishi R, Sankaran S, Grady L, et al. Automated estimation of image quality for coronary computed tomographic angiography using machine learning. *Eur Radiol.* 2018.
7. Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J.* 2017;38(7):500–507.
8. Dey D, Gaur S, Ovrehus KA, et al. Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study. *Eur Radiol.* 2018.
9. Tesche C, De Cecco CN, Baumann S, et al. Coronary CT angiography-derived fractional flow reserve: machine learning algorithm versus computational fluid dynamics modeling. *Radiology.* 2018:171291.
10. Zreik M, Lessmann N, van Hamersvelt RW, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal.* 2018;44:72–85.
11. Dey D, Diaz Zamudio M, Schuhbaeck A, et al. Relationship between quantitative adverse plaque features from coronary computed tomography angiography and downstream impaired myocardial flow reserve by 13N-ammonia positron emission tomography: a pilot study. *Circ Cardiovasc Imaging.* 2015;8(10) e003255.
12. Min JK, Dunning A, Lin FY, et al. Rationale and design of the CONFIRM (COronary CT angiography EvaluatioN for clinical outcomes: an InteRnational multicenter) registry. *J Cardiovasc Comput Tomogr.* 2011;5(2):84–92.

Piotr J. Slomka, PhD, FACC*

Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA

E-mail address: Piotr.Slomka@cshs.org

* Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Ste. A047N, Los Angeles, CA 90048, USA.