Review of: "Implementing Machine Learning to predict the 10-year risk of Cardiovascular Disease"

Rajiv Singla

Potential competing interests: No potential competing interests to declare.

Dear Authors and editor,

Here are few suggestions/comments from medical aspect of this manuscript.

1. Premise of this paper is to compare a ML approach with pre-existing risk scores. For this to be analysed from medical aspect, process of diagnosing a CVD need to be discussed. In a person, risk factors are identified (viz. age, sex, family history, BMI, physical activity, smoking, hypertension, dyslipidemia among others) and in presence of clinical symptoms (like chest pain, breathlessness etc.), further diagnostic tests are conducted (like ECG, Exercise stress test, stress echo, angiography) to ascertain presence of CVD. Now the issue from medical/clinical aspect with this paper is that it compares risk factors based scores to diagnostic based outcomes. This is not fair as information inherent in diagnostic tests is more accurate and specific.
2. Comparing outcome based on this model to outcome from individual diagnostic test like flouroscopy or exercise stress test alone may provide clinically relevant additions to the literature.
3. One point that need to be clarified about original UCI database is: how was presence of CVD (outcome of ML model) was determined in this dataset? From clinical side, it seems to be flouroscopy/angiography itself. Now including that as one of the input parameters has capacity to compromise analysis of this paper.
4. However, I believe this paper is part of evolution of application of ML approach to healthcare data and creates more discussions. This is how technology and healthcare (or people using them) would understand each other better. This “Healthcare-tech divide” can perhaps be sorted by including more practising physicians in tech based research.

Best of luck and regards.