Cardiovascular health: The power of machine learning

Machine learning accurately predicts the association between cardio-respiratory fitness and mortality.

A machine learning approach identified, with 90% accuracy, people at highest risk of death within ten years of conducting a treadmill cardio-respiratory fitness test.

The ability of the circulatory and respiratory systems to supply oxygen to skeletal muscles during sustained physical activity is not only an objective measure of regular physical activity, but also a useful diagnostic and prognostic health indicator. Poor cardio-respiratory fitness has been linked to cardiovascular disease and death from any cause, with many studies highlighting the beneficial effects of exercise.

Statistical modelling is commonly used to predict adverse clinical outcomes based on cardio-respiratory fitness. Previous work has shown that this type of analysis can identify between 30 percent and 50 percent of patients at high risk of death within an eight-year period.

As the use of machine learning to analyse biomedical data expands, Mouaz Al-Mallah at KAIMRC's Clinical Cardiovascular Research Program and colleagues sought to determine whether it can be applied to predict the risk of death within 10 years of undergoing a physician-referred exercise treadmill stress test. "There is a lot of hype about machine learning, but we do not know if it is better than using conventional statistical tools," says Al-Mallah.

The authors used data from 34,212 individuals that took part in the Henry Ford Exercise Testing Project, the largest study of physical fitness to date. The project involved nearly 70,000 adults from Detroit, Michigan, who completed an exercise stress test between 1991 and 2009. Al-Mallah and colleagues focused on patients that were free of coronary artery disease or heart failure at the time of testing and who had a complete 10-year follow-up.

They found that a machine learning approach, involving computer algorithms that can learn from and make predictions from complex data sets, was significantly more accurate than conventional statistical methods at identifying the risk of death in the patient cohort. The machine learning algorithm used data associated with 15 variables — including age, heart rate, blood pressure, reason for the exercise test, and family history — to predict the likelihood of death from any cause. The approach identified, with over 90 percent accuracy, the participants who were at high risk of death.

"Machine learning outperformed statistical tools in every aspect. This is partly due to the large sample size, which is known to be essential for machine learning to perform well," explains Al-Mallah.

The team is now testing machine learning on other outcomes and hopes it will have a similar predictive power. Al-Mallah expects that machine learning will have a huge impact on medical diagnosis and prognosis thanks to its ability to analyse images and recognize patterns. "In the future, the job of the radiographer will be to confirm the machine learning findings rather than to interpret the images," Al-Mallah says.

> "They found that a machine learning approach was significantly more accurate than conventional statistical methods at identifying the risk of death in the patient cohort."

Al-Mallah, M. H., Elshawi, R., Ahmed, A. M., Qureshi, W. T., Brawner, C. A. et al. Using machine learning to define the association between cardiorespiratory fitness and all-cause mortality (from the Henry Ford Exercise Testing Project). *American Journal of Cardiology* **120**, 2078–2084 (2017).