Abstract

Traditional cancer-screening tests are expensive, invasive, and often limited to a single cancer type. A recently developed blood test, CancerSEEK, had shown promise in simultaneously screening for eight different types of cancers: ovarian, liver, stomach, pancreatic, esophageal, colorectum, lung, and breast, of which the first five types previously had no screening tests. However, CancerSEEK's accuracy was limited by its logistic regression model. Presented here are improved results by using a deep learning artificially intelligent neural network capable of exploiting nonlinear relationships between variables that CancerSEEK could not. The new model, called Cancer Detecting Neural Network (CDNN), was fitted to the same dataset as that used in CancerSEEK for fair evaluation comparison. The dataset was generated from a blood test that detected protein biomarker concentrations and gene mutations from each of 1817 patients. CDNN was then evaluated using ten-fold cross validation, and finally optimized by adjusting hyper-parameters and the output threshold. While keeping the specificity above 99%, the median accuracy was increased from 70.80% in the CancerSEEK model to 82.30% using CDNN.