

國立中央大學數學系聯合統計所

專題演講

主 講 人：張憶壽教授 (National Health Research Institutes)

演講題目：From lung cancer risk prediction models to bias-variance tradeoff: some of my experiences in big data

演講茶會：2020年11月24日(星期二) 10:30 a.m. ~ 11:00 a.m.

茶會地點：中央大學鴻經館M306

演講時間：2020年11月24日(星期二) 11:00 a.m. ~ 12:00 p.m.

演講地點：中央大學鴻經館M116

Abstract：

In this talk, I will first discuss some current efforts in early detection of lung cancer and then discuss issues in building clinical prediction models, including performance evaluation criteria, explanation, and bias-variance tradeoff.

Lung cancer is the leading cause of cancer death worldwide. Most lung cancer is diagnosed at late stage and hence prevention and early detection are important. In view of the landmark 2011 National Lung Screening Trial report that a 20% mortality reduction was observed in the low-dose CT (LDCT) screening arm among heavy smokers, USPSTF recommends annual LDCT screening for those ages 55—80, has smoked 30 or more pack-years, and smoked in the past 15 years. In addition, risk prediction models have been proposed to select high risk individuals for LDCT screening for lung cancer in the west. In Taiwan, there is also an enthusiasm in LDCT screening for lung cancer among never-smokers as well as smokers. We proposed a risk prediction model for never-smoking females in this regard and showed that it performs well, compared with the USPSTF smoking criteria for American women. This model uses only age, BMI, COPD, education, family history, and 9 genetic variants for prediction. I will use this example to discuss performance assessment criteria, explanation and bias-variance tradeoff in statistical learning. In particular, I will mention some of the roles of Hoeffding's inequality and VC class in big data and classical large sample theories.

贊助單位：國立中央大學數學系

敬 請 公 告 ， 歡 迎 參 加