

# Yale SCHOOL OF PUBLIC HEALTH

## *Biostatistics*

---

### ***Presentation***

---

Combining Biomarkers Nonlinearly for Classification Using the  
Area Under the ROC Curve

**Shuxin Yin, PhD**

Post Doctoral Research Fellow

Vaccine & Infectious Disease Division

Fred Hutchinson Cancer Research Center

#### **ABSTRACT**

In biomedical studies, it is often of interest to classify/predict a subject's disease status based on some biomarker measurements. Two approaches have received a lot of attention in the biostatistical literature for finding optimal biomarker combinations using a training data. The logistic approach maximizes a logistic regression model likelihood, while the AUC (area under the receiver operating characteristic curve) approach maximizes the empirical AUC based on biomarker combination. I propose a new method for maximizing the empirical AUC that approximates the AUC loss function with a non-convex, non-smooth ramp function. The optimization procedure naturally lends itself to the use of 'kernel trick', which may obtain a better combination of biomarkers when there is strong nonlinearity in the data. I demonstrate through simulation studies and real data analysis that the proposed method not only outperforms existing AUC approach methods in finding the best linear combination of markers in the original space, but it can also successfully capture nonlinear pattern in the data to achieve much better classification performance.

Tuesday March 5, 2013 11:30 am-12:30 pm  
60 College Street LEPH Room 101

---