

# THE 10th ANNUAL ABBOTT LABORATORIES DISTINGUISHED LECTURESHIP IN PHARMACEUTICAL APPLICATIONS

## A Statistics and Biostatistics Joint Seminar

*Speaker:*

Mark van der Laan  
Division of Biostatistics  
University of California - Berkeley

*Title:*

Towards Robust Machine Learning Algorithms for  
Causal Effects that Preserve Meaningful Statistical Inference

*Place & Time:*

Wednesday, April 22, 2009  
Room 140 Bardeen  
4 pm

*Abstract:*

Current statistical practice to assess an effect of an intervention or exposure on an outcome of interest often involves either maximum likelihood estimation for a priori specified regression model, or, manual and/or data adaptive interventions to fine tune a choice of model. In both cases, bias in the point estimates and the estimate of the signal to noise ratio are rampant, causing an epidemic of false claims based on data analyses.

In this talk we present our efforts to construct machine learning algorithms for estimating a causal effect that take away the need for specifying regression models, while still providing maximum likelihood based estimators and inference. Two fundamental concepts underlying this methodology are the very aggressive use of cross-validation to select optimal combinations of many model fits, and subsequent targeted maximum likelihood estimation to target the fit towards the causal effect of interest.

We illustrate this method in observational studies for assessing the effect of an intervention on adherence to drug regimen in HIV infected patients, and for discovery of mutations in the HIV virus that cause resistance to a particular drug regimen.

We also illustrate the performance on FDA approved clinical trials, simulated data imitating postmarket safety analysis, and the analysis of single nucleotide polymorphisms.

Coffee and Cookies at 3:30 p.m. in Room 1210 MSC