

HARVARD UNIVERSITY
DEPARTMENT OF STATISTICS



Tel. 617-495-5496
Fax. 617-496-8057
statdept@stat.harvard.edu
<http://www.stat.harvard.edu/>

Science Center 7th Fl.
One Oxford Street
Cambridge, MA 02138

STATISTICS COLLOQUIUM

MONDAY, APRIL 11, 2016
TALK: 4:15 PM — SCIENCE CENTER 300H
RECEPTION: 3:50 PM

**“Does the Method Matter More than the Truth?
Effects of Method Parameters and Ground Truth
in the OMOP Results Database”**

[Alan F. Karr](#)

Center of Excellence for Complex Data Analysis, RTI International

ABSTRACT

The OMOP results database is a unique dataset containing the results of more than 6 million statistical analyses of patient-level medical adverse event data. The patient data are drawn from five large provider databases, and cover 181 drugs and 19 outcomes. Seven classes of statistical methods were applied to each (drug, outcome) pair, with 1246 variants arising from parameter settings within each method class. The dataset also contains a binary ground truth for each (drug, outcome) pair, drawn from the scientific literature.

This presentation will describe a set of condition-specific analyses of the OMOP results database that illuminate the roles of modeling choices and ground truth in estimating log relative risks. The analyses are partition models that show concretely the effects and interactions of model parameters and ground truth, as well as reveal clearly the bias resulting from particular choices of parameters.

An simplified encapsulation is that how the data are analyzed may have more influence on the results more than does ground truth.