

**Data Science Seminar**  
Hosted by Department of Mathematical Sciences

- Date: Tuesday, April 16, 2019
- Time: 11:45am - 12:45pm
- Room: WH-G02 (special location)
- Speaker: David Madigan (Columbia University)
- Title: Towards honest inference from real-world healthcare data

**Abstract**

In practice, our learning healthcare system relies primarily on observational studies generating one effect estimate at a time using customized study designs with unknown operating characteristics and publishing – or not – one estimate at a time. When we investigate the distribution of estimates that this process has produced, we see clear evidence of its shortcomings, including an apparent over-abundance of estimates where the confidence interval does not include one (i.e. statistically significant effects). We propose a standardized process for performing observational research that can be evaluated, calibrated and applied at scale to generate a more reliable and complete evidence base than previously possible, fostering a truly learning healthcare system. We demonstrate this new paradigm by generating evidence about all pairwise comparisons of treatments for depression for a relevant set of health outcomes using four large US insurance claims databases. In total, we estimate 17,718 hazard ratios, each using a comparative effectiveness study design and propensity score stratification on par with current state-of-the-art, albeit one-off, observational studies. Moreover, the process enables us to employ negative and positive controls to evaluate and calibrate estimates ensuring, for example, that the 95% confidence interval includes the true effect size approximately 95% of time. The result set consistently reflects current established knowledge where known, and its distribution shows no evidence of the faults of the current process.

About the speaker: David Madigan is Professor of Statistics at Columbia University. From 2013 to 2018 he was the Executive Vice-President for Arts and Sciences and Dean of the Faculty of Arts and Sciences at Columbia. He is also a member of the Data Science Initiative at Columbia and part of the OHDSI (the Observational Health Data Sciences and Informatics) program. He received a bachelor's degree in Mathematical Sciences and a Ph.D. in Statistics, both from Trinity College Dublin. He has previously worked for AT&T Inc., Soliloquy Inc., the University of Washington, Rutgers University, and SkillSoft, Inc. He has over 120 publications in such areas as Bayesian statistics, text mining, Monte Carlo methods, pharmacovigilance and probabilistic graphical models. He is an elected Fellow of the American Statistical Association, the Institute of Mathematical Statistics, and the American Association for the Advancement of Science. He recently completed a term as Editor-in-Chief of Statistical Science and is the current editor of Statistical Analysis and Data Mining.

From: <https://www2.math.binghamton.edu/> - **Binghamton University Department of Mathematics and Statistics**

Permanent link: <https://www2.math.binghamton.edu/p/seminars/datasci/190416>

Last update: **2019/04/05 14:57**



